Making ERP research more transparent: Guidelines for preregistration

Mariella Paul\textsuperscript{a,b,c,1}, Gisela H. Govaart\textsuperscript{a,b,d,1}, Antonio Schettino\textsuperscript{e,f,*,1}

\textsuperscript{a} Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany
\textsuperscript{b} Humboldt-Universität zu Berlin, Faculty of Philosophy, Berlin School of Mind and Brain, Luisenstraße 56, 10117 Berlin, Germany
\textsuperscript{c} Psychology of Language Department, University of Göttingen, Gotthelfstraße 14, 37073 Göttingen, Germany
\textsuperscript{d} Charité – Universitätsmedizin Berlin, Einstein Center for Neurosciences Berlin, 10117 Berlin, Germany
\textsuperscript{e} Erasmus Research Services, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, the Netherlands
\textsuperscript{f} Institute for Globally Distributed Open Research and Education (IGDORE), Sweden

\textsuperscript{1} Corresponding author at: Erasmus Research Services, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, the Netherlands.

\textsuperscript{*} Equal contribution.

\textsuperscript{1} E-mail address: schettino@euv.nl (A. Schettino).

ABSTRACT

A combination of confirmation bias, hindsight bias, and pressure to publish may prompt the (unconscious) exploration of various methodological options and reporting only the ones that lead to a (statistically) significant outcome. This undisclosed analytic flexibility is particularly relevant in EEG research, where a myriad of preprocessing and analysis pipelines can be used to extract information from complex multidimensional data. One solution to limit confirmation and hindsight bias by disclosing analytic choices is \textit{preregistration}: researchers write a time-stamped, publicly accessible research plan with hypotheses, data collection plan, and the intended preprocessing and statistical analyses before the start of a research project. In this manuscript, we present an overview of the problems associated with undisclosed analytic flexibility, discuss why and how EEG researchers would benefit from adopting preregistration, provide guidelines and examples on how to preregister data preprocessing and analysis steps in typical ERP studies, and conclude by discussing possibilities and limitations of this open science practice.

1. Introduction

Over the last decade, findings from a number of research disciplines have been under careful scrutiny. Prominent examples of research supporting incredible conclusions (Bem, 2011), failures to replicate popular and highly cited published findings (Board of Governors of the Federal Reserve System et al., 2015; Camerer et al., 2016; Errington et al., 2014; Open Science Collaboration, 2015), sloppy scientific practices (van der Zee et al., 2017), and breaches of ethical conduct (Levelt et al., 2012) increased the suspicion that published results might be inflated or incorrect (Goldacre et al., 2019; Hannink et al., 2013; Ioannidis, 2008, 2005; Jones et al., 2017; Simmons et al., 2011; Trinquart et al., 2018), resulting in considerable waste of resources (Chalmers et al., 2014) and, at times, life-threatening consequences (Anand et al., 2014; Topol, 2004; Vedula et al., 2012). These events motivated mainstream discussions on incentive structures (Edwards and Roy, 2017; Nosek et al., 2012), statistical literacy (Cumming, 2014; Kruschke and Liddell, 2017; Wasserstein and Lazar, 2016), and theoretical and methodological rigor (Devezer et al., 2020; Eronen and Bringmann, 2021; Oberauer and Lewandowsky, 2019; Szollosi and Donkin, 2021). At the heart of all these proposed reforms lies a call for increased transparency in scientific reporting (Nosek and Bar-Anan, 2012; Simmons et al., 2012; Simonsohn, 2013; Wilson et al., 2017). Transparency at all research stages effectively mitigates confirmation bias – searching, interpreting, and remembering information that supports prior beliefs while ignoring evidence against them (Nickerson, 1998) – and hindsight bias – the tendency to overestimate the extent to which past events were able to predict a present outcome (Roese and Vohs, 2012).

These cognitive biases find fertile ground in complex and multifaceted intellectual endeavors like empirical sciences. Data collected in an experimental or observational study are rarely interpreted in their raw form. Instead, researchers typically apply a series of transformations to deal with outliers and missing data (Enders, 2010; Hawkins, 1980), combine or discretize variables into composite indices, change the unit of measurement, and so on. In other words, “data are to a certain extent actively constructed” (Steegen et al., 2016, p. 702). Moreover, there are countless statistical techniques that can be chosen to analyze the preprocessed data, including classical null hypothesis tests (Field et al.,...
2 Readers may be familiar with other terms, such as researcher degrees of freedom (Simmons et al., 2011) or garden of forking paths (Gelman and Loken, 2013), which refer to all choices that researchers make throughout their workflow, including hardware and software selected for data collection and analysis, the type and number of stimuli presented to participants, and much more. These choices can have tangible consequences on study results and interpretation. For instance, the same analysis pipeline on the same dataset can lead to quantifiably different results when run with different software (Wood, 2019; see also Eklund et al., 2016). Typically, larger and more homogeneous samples of both participants and stimuli increase statistical power, which in turn can lead to different interpretations of the results. Quoting Sandre et al. (2020): “[…] different ways of processing the same data can lead researchers to different conclusions, demonstrating yet again that transparency of all processing decisions is a necessity.” (p. 35). We concur: transparently reporting all analytic choices would increase study reproducibility and, more generally, the trustworthiness of the electrophysiological literature.

As mentioned before, researchers (just like other human beings) tend to shape their analytic choices with the (largely implicit) aim to confirm their prior beliefs, and post-hoc justification of said choices is rationalized under the “illusion of objectivity” (Kunda, 1990; Pyszczynski and Greenberg, 1987). EEG researchers are not immune to the pitfalls of confirmation bias and hindsight bias: for example, they may be thinking that all preprocessing and analysis choices were determined a priori while they may have been at least partly based on seeing the data. Sandre et al. (2020) suggest that “a single processing stream should be
finalized before any analyses are undertaken" (p. 35). In other words, confirmation bias and hindsight bias cannot take place if analytic choices are not only determined before the data are collected but also transparently reported. This practice is called preregistration.

2. Preregistration

Preregistrations are time-stamped, (eventually) publicly accessible documents with hypotheses, data collection plan, and/or intended preprocessing and statistical analyses, written before the start of a research project. In other words, researchers commit to one among many ways in which the study can be conducted and analyzed. This document is uploaded on a trusted online repository – e.g., Open Science Framework (OSF; https://osf.io/), ClinicalTrials.gov (https://clinicaltrials.gov/), American Economic Association’s registry for randomized controlled trial (AEA RCT Registry; https://www.socialscienceregistry.org/) –, which assigns it a date and time. The protocol is made public immediately or after an embargo period. Date and time of submission ensure that the research plan was devised before starting the study.

The popularity of preregistration has skyrocketed in recent years: for example, the number of publicly available preregistrations uploaded on the OSF went from 38 in 2012 to 36,675 in 2019 (Bakker et al., 2020). Many journals now explicitly encourage this practice by awarding “preregistration badges” (https://osf.io/vyyzx/; see Kidwell et al., 2016), including Psychological Science and Cortex (for a full list, see https://tinyurl.com/COS-badges).

2.1. Advantages of preregistration

The advantages of preregistering research plans are manifold. First and foremost, preregistration can be seen as an additional tool to effectively achieve as much transparency as possible (see also Navarro, 2020), ultimately increasing verifiability at all stages of the research cycle (Resnik, 2005; Lupia and Elman, 2014; see also Merton, 1942). Researchers are expected to abide by ethical principles that are functional to the epistemic goals of science: advancing human knowledge by describing nature, developing theories and hypotheses that allow the generation of reliable predictions, and eliminating errors and biases (Resnik, 2005). Openness is one of these foundational principles: “Scientists should share data, results, methods, ideas, techniques, and tools. They should allow other scientists to review their work and be open to criticism and new ideas.” (Resnik, 2005, p. 52). Other open science practices – e.g., sharing study protocols, materials, raw data, and analysis code – directly follow from this principle. Preregistration additionally offers the possibility to document the rationale behind theoretical and methodological choices, useful not only in quantitative but also qualitative disciplines (Haven and Van Grootel, 2019). In addition, deviations from the original design (e.g., discrepancies between planned and actual sample size, unforeseen moderators, flexible exclusion criteria) can be more easily identified, effectively counteracting selective outcome reporting (Goldacre et al., 2019; John et al., 2012; Simmons et al., 2011). Furthermore, the presence of public, traceable evidence of the original plan exposes (and possibly mitigates) confirmation bias and hindsight bias.

Preregistration allows researchers to specify the rationale and hypotheses of the study while also maintaining flexibility with respect to additional analyses conducted after seeing the data, provided that they are included in a different section of the final manuscript. This precludes presenting any hypotheses generated after observing the data as if they were a priori, or “hypothesizing after the results are known” (HARKing; Kerr, 1998). This practice is particularly difficult to identify in published papers because readers can only access what the authors reported after collecting, analyzing, and interpreting the data, without knowing whether the hypotheses described in the introduction were originally unanticipated (or even considered implausible) until reassessed in light of the collected empirical evidence. This problem is magnified by the fact that, at least in some research fields, theoretical frameworks and hypotheses are often underspecified, which decreases their explanatory power and predictive utility (Meehl, 1967; Muthukrishna and Henrich, 2019; Szollosi and Donkin, 2021; van Rooij and Baggio, 2021).

It has also been argued (Nosek et al., 2019, 2018) that preregistration can contribute to mitigating publication bias in the academic literature (Nissen et al., 2016; Rosenthal, 1979; Scargle, 2000), since research plans are discoverable regardless of whether the final report is ultimately published in peer-reviewed journals. Yet, in our opinion, publication bias can only be effectively mitigated when results are published regardless of study outcome, that is, via Registered Reports (see Section 4.1) or journals that publish studies based on scientific rigor rather than their outcome (see Section 4.2). Nonetheless, discoverability of research plans is a useful step in making the entire research process discoverable. Importantly, preregistration is not only helpful when hypotheses are tested or p-values are reported (McPhetres, 2020), but also for exploratory and qualitative research (Dirnagl, 2020; Haven and Van Grootel, 2019) and when using other statistical procedures (e.g., specify and justify in advance what priors will be used in a planned Bayesian analysis; see Depaoli and van de Schoot, 2017).

Correlational evidence accumulated over the past 20 years in several disciplines suggests that preregistration may facilitate the publication of non-significant findings, thus providing a more accurate representation of available knowledge. For instance, Kaplan and Irvin (2015) reviewed a sample of randomized clinical trials funded by the National Heart, Lung, and Blood Institute evaluating drugs or dietary supplements for the treatment or prevention of cardiovascular disease. Of the 55 selected studies, 30 were published before and 25 after the year 2000, when study registration on ClinicalTrials.gov became compulsory in the U.S. following the Food and Drug Administration Modernization Act in 1997. Results showed that 57% of the studies published before 2000 showed a significant benefit of the intervention, as opposed to only 8% of trials published after 2000. Similar results were reported in a (preregistered) meta-analysis of meta-analyses of orthodontics and dentofacial orthopedics studies: registered trials reported less favorable intervention effects compared to unregistered trials (Papageorgiou et al., 2018). Preregistration can also help identify whether funding sources are correlated with study outcome, potentially uncovering questionable practices due to (undisclosed) conflicts of interest. For instance, a review of studies of safety and efficacy trials for a wide array of drugs (Bourgeois et al., 2010) revealed that trials funded by industry were less likely to be published within 2 years from study completion and most likely to report a positive outcome (85%, as opposed to 50% for government-funded trials).

---

It should be noted that the results of statistical tests can still be valid (i.e., expected false positives close to nominal α) assuming proper statistical conditioning, e.g., by building the conditional reference distribution of the test statistic via data permutation (for details, see Devezer et al., 2020, Box 2).

A preregistration can serve as a ‘log’ for exploratory research, to make the many choices during the research process transparent: “Methodological and analytic flexibility is maintained but disclosed.” (Dirnagl, 2020, p. 4).

The clinical studies considered in Kaplan and Irvin (2015) were specifically chosen to be large, well-funded projects, likely to get published even if results were not statistically significant. Thus, their work does not directly show that preregistered studies are easier to get published. Yet, it does suggest that, if studies are preregistered, non-significant findings are more likely to be reported as such, instead of being p-hacked to chase publication.
Recently, Adda et al. (2020) analyzed the distribution of p-values of primary outcomes for phase II and phase III drug trials registered on ClinicalTrials.gov between 2010 and 2019 and found no indication of selective outcome reporting, suggesting that such registries may successfully disincentivize the (conscious or unconscious) use of suboptimal reporting practices and, consequently, improve the credibility of published research.

2.2. Benefits for individual researchers

Besides being advantageous for whole research fields, anecdotal experience and preliminary evidence suggest that preregistration can be beneficial for individual researchers as well (Allen and Mehl, 2019; Mekierman et al., 2016; Toth et al., 2020; Wagenmakers and Dutilh, 2016). Generally speaking, drafting a thorough preregistration – preferably with the help of useful templates and checklists (see Section 3.1) – can improve the experimental design not only because authors are stimulated to think more carefully about the research plan, but also because feedback from peers can be solicited early and incorporated when most valuable, that is, when there is still time to make changes. Early-career researchers (ECRs) may benefit even more from learning this skill, since they are often directly involved with the ideation and development of the research project, data preprocessing and analysis, and writing of the final report. Preregistering a study as an ECR can also give a stronger sense of ownership over ideas that were originally conceptualized by their supervisors, for example by having a clearer overview on the different steps of the workflow, making informed decisions about the rationale, experimental design, and planned analyses early on in the project.

Given an increasing interest in transparency, we expect ECRs to be working in an environment that values – and might even require – a certain level of commitment to open science practices, of which preregistration is an example. As mentioned earlier, a growing number of journals encourage preregistration, e.g., by means of badges. In academia, funding agencies appreciate the importance of study preregistration in medical and non-medical disciplines: for example, the recent COVID-19 Programme by the Dutch funder ZonMw includes specific open science guidelines for prospective applicants, among which mandatory preregistration of animal studies and “strongly recommended” preregistration for all other studies. In addition, preregistration may lead editors and reviewers to more easily trust authors when reporting certain methodological choices, such as sequential testing and one-sided tests (Lakens, 2017, Study 1). Last but not least, researchers who preregister their studies may be perceived as more trustworthy, because they are willing to open all products of their workflow to their peers for scrutiny. However, a recent registered report investigating whether preregistration increases peers’ trust in the final publication revealed inconclusive evidence either in favor or against this hypothesis (Field et al., 2020), leaving this question open for future examinations.

The advantages of preregistration in neuroimaging and electrophysiology have not yet systematically been evaluated. Nonetheless, the data accumulated in other disciplines provide a number of insights, practical examples, and learned lessons that can guide a widespread and informed implementation of this practice in our research field. When done properly, preregistration works as intended.

3. Recommendations for preregistration of ERP research

In this section, we provide guidelines on how to transparently document the planned analytic choices in a preregistration of a prototypical ERP study. We focus on ERPs because of their widespread use in cognitive and clinical research (Hajcak et al., 2019; Helfrich and Knight, 2019), although most of these recommendations can still be useful when using other signal processing techniques (e.g., ERP and time-frequency analyses have many preprocessing steps in common). Furthermore, we only include sections that would decrease researchers’ flexibility during signal preprocessing and statistical analysis. As discussed earlier, these steps are complex and multifaceted, with many reasonable choices that can lead to qualitatively different interpretations of the data; therefore, transparently documenting them would have a significant impact on the verifiability of the results. However, other aspects of a study should also be carefully planned and included in the preregistration protocol, e.g., the rationale behind the chosen sample size (including a power analysis; for recent guidelines, see Baker et al., 2020; Boudewyn et al., 2018), inclusion and exclusion criteria, and stimulus details and characteristics (e.g., to ensure that items sampled from all planned conditions are reported in the published manuscript).

We encourage researchers to craft a document that is specific, precise, and exhaustive (Veldkamp, 2017, chap. 6; Wicherts et al., 2016). A preregistration is specific when it includes a detailed description of all phases of the research workflow, from the initial design of the study to the information reported in the final manuscript; precise when the research plan is interpretable in only one way (e.g., there is no ambiguity regarding the intended preprocessing pipeline); and exhaustive when the research plan states that only the mentioned analyses will be considered diagnostic to confirm or falsify predictions, thereby clarifying that other analyses have been conducted after seeing the data (see also McPhetres, 2020, on adding underspecified secondary analyses).

In our experience, it is very useful to run a pilot study before drafting the preregistration document. Advantages include: (i) gauge the feasibility of recruitment, randomization, and assessment procedures, especially when testing clinical populations and/or evaluating a novel treatment (e.g., Leon et al., 2011); (ii) ensure that task instructions are clear for participants; (iii) confirm that the target ERP component(s) are elicited; (iv) test preprocessing and analysis pipelines for possible bugs, errors, and/or computational feasibility. We emphasize that small scale pilot studies should not be used to estimate effect sizes to inform a priori power analysis (for details, see Albers and Lakens, 2018; Kraemer et al., 2006).

3.1. Preregistration templates

In principle, any time-stamped, accessible protocol with a clear study plan can serve as a preregistration. However, ready-made templates can greatly facilitate the inclusion of preregistration in researchers’ workflows by providing a list of bullet points (Bakker et al., 2020; Wicherts et al., 2016). In addition, hosting preregistrations on online platforms that are popular among the research community (rather than, for example, personal websites) can improve accessibility. One of the most popular platforms is the OSF, which offers several preregistration templates differing on topic, length, and specificity. While less extensive templates (e.g., AsPredicted: https://aspredicted.org/) are typically used by newcomers for their first preregistration, we would rather recommend the standard OSF Prereg template, whose increased level of detail facilitates the creation of specific, precise, and exhaustive preregistrations that more effectively decrease the risk of undisclosed analytic flexibility. A template specifically for preregistration of EEG studies was started during a hackathon at the annual meeting of the Society for the Improvement of Psychological Science (SIPS) in 2019.
(Algermissen et al., 2019) and is currently being developed online by an active community of volunteers. Readers are welcome to contribute to (and use) the current draft at https://tinyurl.com/eegprereg.

Below we provide some examples on how to preregister typical preprocessing and analysis steps in an ERP study. We do not intend to recommend one preprocessing step or statistical method over another, but rather give examples on how commonly used preprocessing and analysis steps can be transparently reported. Please note that a good preregistration should also be explicit about the order of preprocessing steps. Again, the examples below are for illustrative purposes: their order is not meant to be prescriptive, and should be adjusted based on the pipeline that is appropriate for the specific study.

It can also be advantageous to include analysis scripts in the preregistration (see, for instance, Nunez et al., 2017). In any case, the preregistration should be specific about the software and (standardized) pipelines that will be used to carry out the preprocessing steps. If researchers plan to use the default settings of a given software, they should also include its version number (these settings might change with different versions) and clearly state in the preregistration that default parameters will be used.

3.2. Preprocessing

3.2.1. Re-referencing

A preregistration should specify which electrodes will be chosen for the offline re-referencing. Common offline reference channels include the linked mastoids, ears, vertex, or an average reference. An example could be: “The continuous data is [...] re-referenced to the average of the left and right mastoid.” (preregistration: Nieuwland et al., 2018a; publication: Fleur et al., 2020).

3.2.2. Filtering

Preregistering the parameter values of the filters that will be applied to the recorded EEG data should be detailed enough to theoretically allow readers to completely reproduce each filter (see Widmann et al., 2015). This includes specifying not only the filter cut-off frequency, but also the type (e.g., butterworth, finite impulse response, infinite impulse response), transition width, passband edge, the order for the transition bandwidth, at what point during the preprocessing pipeline the filter was applied (e.g., to continuous or segmented data), and – in case of multiple filters – the order in which the filters are applied. For example: “The continuous EEG data will be filtered with separate Hamming windowed sinc finite impulse response (FIR) filters (Widmann, 2006): (1) high-pass: passband edge 0.5 Hz, filter order 1,690, transition bandwidth 0.5 Hz, cutoff frequency (−6 dB) 0.25 Hz; (2) low-pass: passband edge 30 Hz, filter order 114, transition bandwidth 7.4 Hz, cutoff frequency (−6 dB) 33.71 Hz.” (preregistration: Schettino et al., 2017; publication: Schindler et al., 2018).

3.2.3. Trial segmentation and time-locking

For trial segmentation, it is especially important to specify when the continuous EEG data are segmented, because this has implications for other preprocessing steps. When preregistering trial segmentation, it is also important to specify what the trial will be time-locked to – e.g., stimulus onset, participant’s motor response – and how long the pre- and post-stimulus period will be. For example: “For the ERPs, we [...] epoch the data from -500 to 1500 ms relative to [critical word] onset.” (preregistration: Coopmans and Nieuwland, 2018, publication: Coopmans and Nieuwland, 2020).

3.2.4. Interpolation

With interpolation, the EEG signal recorded from noisy channels is replaced with estimated activity from neighboring electrodes. In the preregistration protocol, one should preregister the criteria used to identify noisy channels as well as the algorithm that will be used for the interpolation. For example: “Bad channels with a voltage > 2 SD of the EEG voltage will be interpolated with a spline interpolation [...]”. The interpolation algorithm [...] is implemented in the EEGLAB toolbox.” (preregistration: Duma et al., 2018, publication: Duma et al., 2019).

3.2.5. Artifact rejection and correction

Approaches to artifact rejection can be roughly divided into three categories: (1) automatic; (2) manual; and (3) semi-automatic, that is, a combination of automatic and manual approaches. Many different algorithms for automatic artifact rejection are available and implemented in different software and toolboxes. For example, the FASTER algorithm (Nolan et al., 2010) calculates various statistical parameters of the signal and defines data as artifactual when exceeding a pre-specified z-score (e.g., ±3). This approach can be additionally integrated by identifying the frequency bands in which one would expect artifacts to occur, e.g., 110 to 140 Hz for muscle artifacts (see Delorme et al., 2007). From a computational perspective, fully automatic approaches are more reproducible, although their sensitivity and specificity can vary. Semi-automatic and manual approaches are more subjective and dependent on the researcher’s skills, but they may be necessary when working with special populations (e.g., infants) whose signal shows less typical artifacts, difficult for automatic approaches to detect. Therefore, to increase transparency and reproducibility, the final report should include the list of epochs marked for rejection (or correction; see below). In addition, source code could be referenced in-text and be made publicly available for inspection upon publication, to provide additional information (e.g., on some parameter values) while decluttering the Methods section. An example of a planned automatic artifact rejection procedure pertaining specifically to ocular artifact rejection could read: “All trials will be checked for eye and muscle activity related artifacts. To detect eye movements and blinks, the EOG signals will be combined to derive bipolar vertical and horizontal channels that will be passed through a set of artifact detection steps. Any trials containing amplitude change larger than 80 µV in the vertical bipolar EOG channel and 120 µV in any other channel within a moving window of 200 ms will be removed to avoid any contamination of data by eye blinks or muscle movements. Additionally, any trials with potential eye-movement activity, i.e., amplitude changes larger than 25 µV in the horizontal bipolar EOG channel detected by a step function, will also be rejected. Participants with rejection rates larger than 30% of the total trials in any of the experimental conditions will be excluded from further analyses.” (preregistration: Bocinova and Johnson, 2017, publication: Bocinova and Johnson, 2019).

In addition to artifact rejection procedures, there are a variety of techniques available to correct artifacts in order to reduce data loss. Methods for artifact correction include independent component analysis (ICA), regression-based methods, and wavelet-transforms (see Jiang et al., 2019). Each of these methods require different parameter values to be preregistered. For example, for ICA, at least the following information should be added (see, for instance, Keil et al., 2014): the method used to compute the ICA – e.g., fastICA (Hyvärinen, 1999) or infomax ICA (Bell and Sejnowski, 1997) –, the electrodes included, the number of

12 Other examples of preregistration can be retrieved from a public Zotero Group library maintained by the Center for Open Science (https://tinyurl.com/OSF-Zotero-group) as well as a spreadsheet started during the aforementioned hackathon at SIPS2019 (https://tinyurl.com/SIPS2019-prereg-list).

13 As already mentioned, high-pass filters may produce edge artifacts when applied to non-continuous signal; thus, segmentation is typically performed after these filters are applied. Other methods are also effective, e.g., extending epochs with zeroes (zero-padding) so that edge artifacts do not affect the actual signal.

14 Recently, though, automatic artifact rejection pipelines for infant EEG data have been developed, e.g., MADE (Debnath et al., 2020) and HAPPE (Gabard-Durnam et al., 2018).
computed components, the method by which artifactual components will be identified – e.g., using templates (Campos Viola et al., 2009) or manually –, and which type of artifactual components will be removed (e.g., only ocular components, ocular and heart beat components, etc.). For example, a preregistration for artifact correction using ICA could read: “Data are then subjected to independent component analysis using single-order blind identification [...]. This is achieved by transforming the weight matrix for components into z-scores across all electrodes, and identifying those that have a z-score greater than 4.0. This is an arbitrary large value which has been determined in previous studies to identify signals due to blinks or to other artefact. Components whose activity is heavily focused on a single electrode are then subtracted from the signal.” (preregistration: Hobson and Bishop, 2014; publication: Hobson and Bishop, 2016).

3.2.6. Baseline correction

A preregistration should also clarify whether or not the data will be baseline-corrected by setting the scalp distribution to zero during a preset period before the onset of the event of interest. Important information include the time window that is used as the baseline, and at which point in the analysis pipeline the baseline correction will be applied. For example: “Epochs extending from −200 ms to +1000 ms time-locked to word onset will be created, and baseline correction will be applied using the pre-stimulus interval.” (preregistration: Schettino et al., 2017; publication: Schindler et al., 2018).

3.3. Statistical analysis

Analytic flexibility does not only occur at the level of preprocessing. Rich M/EEG data sets can be analyzed in multiple different ways – e.g., event-related potentials (ERPs; Luck, 2014), EEG microstates (Michel and Koenig, 2016), time-frequency (Cohen, 2014), functional connectivity (Bastos and Schoffelen, 2016), steady-state evoked potentials (Regan, 1977), source localization (Michel and He, 2019) –, with different objectives, dependent variables, and levels of analytic sophistication. Moreover, researchers have multiple valid options for statistical analysis, such as ANOVAs and t-tests, as well as cluster-based permutation tests (Maris and Oostenveld, 2007), Bayes factors and linear mixed effects models (Florin et al., 2018), threshold-free cluster enhancement (Molenberghs and Khatami, 2013), and more. An example of ERP experiment may involve measurements from 64 electrodes and a sampling rate of 256 Hz, with a trial length of 1,000 ms after stimulus onset (e.g., Schindler et al., 2018). After averaging over trials, this would result in 16,384 data points for each participant and condition. For statistical analysis, this leads to a large number of potential comparisons, often referred to as the multiple comparison problem (MCP). Standard statistical correction procedures operating at the level of single electrode-time pairs would yield hyper-conservative results (increased Type II error); on the other hand, failing to correct for multiple comparisons can easily lead to spurious statistically significant results (increased Type I error). Therefore, statistical plans of ERP studies should always include a strategy on how to deal with the MCP. Several solutions are available (Luck, 2014, chap. 10; Luck and Gasperlin, 2017), including: (i) a priori definition of electrode sites and time windows based on previous studies; (ii) collapsed localizers, i.e., averaging all trials of all conditions of all participants to identify electrode clusters and time windows, thereby avoiding condition-specific biases; and (iii) mass univariate statistics, i.e., computing statistical tests at each electrode and time point and applying appropriate multiple comparisons correction techniques (Fields and Kuperberg, 2020; Groppe et al., 2011).

In the following sections, we provide examples on how to preregister three common ERP analyses: (1) Analysis of Variance (ANOVA); (2) cluster-based permutation tests; and (3) Bayes factors. We hope these examples will be useful for readers planning to preregister similar statistical analyses and give an idea of the preferred level of detail for other analysis methods not listed here.

3.3.1. ANOVA

ANOVA are a popular statistical technique to analyze ERP data: it is not uncommon to read published studies including several within- and between-subject factors for various spatial and temporal ROIs. However, a growing number of factors comes at a cost, namely an increase in false positive rate: up to 50% chance to find at least one false positive effect with 4 factors, and up to 100% chance with 8 factors (Luck and Gasperlin, 2017). Therefore, ERP researchers should carefully plan appropriate corrections for multiple testing not only for the follow-up tests to an ANOVA, but also as a function of the number of factors included in the ANOVA itself. Alternatively, researchers can limit the number of factors included in the ANOVA if they have a specific hypothesis about the spatial and temporal region of interest or only run planned (paired) contrasts on the relevant comparisons. Preregistration can help make these decisions beforehand, without being biased by seeing the data.

In the following example, researchers interested in the N400 ERP component (Kutas and Hillyard, 1980; Kutas and Federmeier, 2011; see also the work by Nieuwland and colleagues for a series of preregistered N400 studies, e.g., Nieuwland et al., 2018b; Coopmans and Nieuwland, 2020) identify the time window and region of interest a priori, and aim to test two hypotheses: (1) larger N400 component for semantically incongruent compared to semantically congruent sentences (i.e., an incongruency effect); and (2) larger incongruency effect for native speakers compared to non-native speakers. In the Analysis section, the preregistration could read: “We will analyze the amplitude15 of the N400 by means of a mixed ANOVA with 2 factors: congruency (semantically congruent sentences vs. semantic violations; within-subject) and language experience (native speaker vs. non-native speaker; between-subject).” As outlined above, researchers should also consider correcting for multiple comparisons as a function of the number of factors in the ANOVA. In this example, three p-values are computed: one for the main effect of congruency, one for the main effect of language experience, and one for their interaction. Researchers could preregister the correction in the following way: “The significance level for the main effect and interaction terms will be Bonferroni-corrected for the number of tests computed in the ANOVA: 0.05/3 = 0.0167.” Of course, the alpha level of the test should also be explicit (α = 0.05 in this example). If researchers also plan follow-up comparisons, they could add: “In case the interaction between congruency and language experience is statistically significant, we will compute two paired t-tests comparing congruent vs incongruent sentences, separately for native speakers and non-native speakers. For these two planned comparisons, we will set α = 0.025 (Bonferroni-corrected for two planned tests with an uncorrected alpha of 0.05).”

3.3.2. Cluster-based permutation tests

Cluster-based permutation tests (CBPT; Maris and Oostenveld, 2007) are another popular statistical approach to analyze EEG data. In the following example, researchers have a hypothesis about a difference in ERP responses between two conditions, but not about specific electrodes and time points. The preregistration should include whether the test is

---

15 Other factors, including context-dependent psychometric properties of brain measures (e.g., Clayson and Miller, 2017) and suboptimal reporting of methodological details in published papers (e.g., Clayson et al., 2019; Larson and Carbine, 2017), further complicate the picture and warrant caution on the reliability of the psychophysiological literature (Baldwin, 2017).

16 It is also important to clearly specify the time window and electrode cluster from which the ERP component is scored. This information should be included in the section Measured variables of the preregistration protocol, e.g.: “We will analyze the mean amplitude value in a time window from 300 to 500 ms after stimulus onset, averaged across a cluster of centro-parietal electrodes (C3, C5, CP3, P3, Pz, P4, P7, and P8).”
one-tailed or two-tailed, within-subject, between-subject, or mixed, as well as the alpha level. In addition, there are several parameters that are more specific to the CBPT, such as the alpha at the cluster-level, the method for computing cluster statistics, the minimum number of electrodes that can form a cluster, how neighboring relations between electrodes will be computed, and the number of randomizations. For example, a preregistration using a CBPT could read: “To test within-subject differences between congruent and incongruent sentences, we will compute a two-tailed cluster-based permutation test using ‘ft_statfun_indepsamplesT’ in Fieldtrip (Oostenveld et al., 2011) with α = 0.025 for each tail (i.e., the overall alpha is 0.05). The alpha at the cluster-level will be set at 0.05. Cluster statistics will be computed with a ’maxsum’ approach and clusters will require a minimum of two neighboring electrodes. Neighboring electrodes will be defined via the ‘triangulation’ method implemented in Fieldtrip. Like on the test-level, the clusters will be tested with a two-tailed statistic. One thousand randomizations will be computed via MonteCarlo method to estimate the p-value under the permutation distribution.”

3.3.3. Bayes factors

Problems inherent in accepting the null hypothesis with classical frequentist procedures (e.g., Wagenmakers, 2007) and common misinterpretations of p-values (Colquhoun, 2017; Wasserstein and Lazar, 2016) are leading an increasing number of researchers to explore Bayesian approaches (Etz and Vandekerckhove, 2018; Kruschke and Liddell, 2017). Bayesian inference allows to incorporate prior knowledge into statistical tests, quantify evidence in favor of the null hypothesis – thus discriminating between “absence of evidence” and “evidence of absence” (e.g., Keysers et al., 2020) –, and monitor the evidence as the data accumulate (Rouder, 2014; but see de Heide and Grünwald, 2020). In particular, Bayes factors – “the extent to which the data sway our relative belief from one hypothesis to the other” (Etz and Vandekerckhove, 2018, p. 10) – have gained considerable popularity, also thanks to the development of user-friendly software that facilitate their calculation (e.g., JASP; https://jasp-stats.org/).

Researchers planning to analyze their data using Bayes factors should clarify the software and procedure used for the estimation, a description of the prior specification (i.e., the type of distribution and its parameter values), and an assessment of the robustness of the results under different prior specifications (see also van Doorn et al., 2020). A preregistered description of planned comparisons using Bayes factors could read as follows: “We will analyze the amplitude values of the N1 ERP component using Bayes Factors (BFs; Kass and Raftery, 1995). Two-tailed Bayesian t-tests (Rouder et al., 2009) will be calculated to estimate the degree of evidence in favor of a model assuming differences between conditions relative to a model assuming no differences. The null hypothesis will be specified as a point-null prior (Dirac distribution, standardized effect size δ = 0), whereas the alternative hypothesis will be defined as a Jeffreys-Zellner-Siow (JZS) prior, a folded Cauchy distribution centered around δ = 0 with scaling factors of r = 1, r = 0.707, and r = 0.5, to verify the robustness of the results as a function of changes in the prior (Schönbrodt et al., 2017). Participants will be included as random factors, and their variance considered nuisance. The threshold to identify the winning model is set at BF ≥ 10 or BF ≤ 0.1, typically considered “strong” evidence in favor of the model in the numerator or denominator, respectively (Kass and Raftery, 1995). BFs will be estimated via the R package BayesFactor v0.9.12-2 (Morey et al., 2015) using Markov Chain Monte Carlo sampling (10,000 iterations).” (preregistration: Schettino et al., 2017; publication: Schindler et al., 2018).

4. General considerations

In Section 2 we clarified how preregistration can mitigate some of the issues related to undisclosed analytic flexibility. In Section 2 we provided guidelines and examples on how to preregister common preprocessing and statistical analysis steps in ERP studies. In Section 4 we discuss several considerations that EEG researchers may want to take into account when critically evaluating whether to preregister their studies.

4.1. Preregistration vs. Registered Reports

Throughout this manuscript we have described unreviewed preregistrations (see van’t Veer and Giner-Sorolla, 2016), that is, the protocols uploaded on public repositories are not formally peer-reviewed. A preregistered study can still be rejected by scientific journals for a number of reasons – for example, lack of interest in “negative” (non-significant) findings (Fanelli, 2010), unprofessional peer-review (Gerwing et al., 2020), or submission in the “wrong” day of the week (Boja et al., 2018) –, thus limiting its discoverability (although the experimental protocol may still be publicly accessible). Conversely, reviewed preregistrations – commonly referred to as Registered Reports (Chambers and Tzavela, 2020) – are alternative article formats in which the study proposal is peer-reviewed and conditionally accepted for publication (in-principle acceptance, or IPA), provided that the original plan is followed and deviations are properly documented. Publication is thus independent from study outcome. Preliminary research has shown that this format seems to effectively mitigate publication bias and reduce the prevalence of selective outcome reporting (Scheel et al., 2020; see also Wiseman et al., 2019). Therefore, we consider Registered Reports the state-of-the-art article format for confirmatory research and recommend them over preregistrations. At the time of writing, more than 250 journals from various scholarly disciplines offer Registered Reports alongside traditional submissions, including Psychophysiology (Kell et al., 2020) and the International Journal of Psychophysiology (Larson, 2016). Please note that, while the current manuscript focuses on preregistrations, our recommendations also hold for Registered Reports.

Despite these desirable properties, researchers should take into account the relatively strict submission criteria (e.g., expected statistical power of 90% or higher) as well as the time necessary to review the study plan (typically 2–4 months; Chambers, 2020), during which the project cannot start. For these reasons, preregistration can be seen as an easier, entry-level practice that is advantageous in itself and helps researchers familiarize with the steps required for a future (recommended) Registered Report submission.

4.2. Potential disadvantages of preregistration

Crafting a comprehensive preregistration protocol requires time. A recent survey (Toth et al., 2020) revealed that respondents with previous experience in preregistration invested, on average, around 4 hours to create the initial draft. Our experience suggests that this could be considered a lower bound: the multidimensional nature of EEG data, coupled with the high level of specificity recommended to effectively avoid selective reporting, requires documenting a large number of preprocessing and analysis steps (see Section 3), planned sample size (with a priori power analysis), sampling strategy, inclusion and exclusion criteria, and more. As Allen and Mehler (2019) point out, planning analyses based on existing data is likely easier, because a preregistration involves anticipating possible outcomes, for example in a decision tree, that depend on seeing the data (e.g., whether the assumptions of the planned statistical model are fulfilled). These authors further suggest

---

17 Due to heterogeneous journal policies, IPA protocols may not always be publicly available or easily verifiable (Hardwicke and Ioannidis, 2018), which sometimes makes it difficult for readers to compare registered plans and published papers. Recent developments have tackled the lack of transparency and standardized protocol registration by updating recommended editorial policy templates (Chambers and Mellor, 2018), but we advise to read the target journal’s specific guidelines before submission.

18 The updated list of journals offering Registered Reports can be found at https://cos.io/rr/.

19 As an example, see Registered Reports Submission Guidelines at Cortex (https://tinyurl.com/RR-Cortex).
that implementing open science practices (including preregistration) during a project increases its duration (“In our experience, these additional requirements can easily double the duration of a project.”; p. 4), which might be especially difficult for ECRs on short-term contracts. However, to the best of our knowledge, this is based only on anecdotal evidence and preregistration could also, at least for some projects, save some time.20 We also note that most of the information included in a comprehensive preregistration is also required in the final manuscript, not only to facilitate communication between the authors and other relevant parties (editors, reviewers, and readers) but also to ensure that the methods leading to the conclusions advertised in the paper are reproducible. Publication guidelines for M/EEG studies (Gross et al., 2013; Keil et al., 2014; Pernet et al., 2018) emphasize the need to describe equipment, study materials, preprocessing steps, dependent variables, and analysis pipelines, and also provide a checklist that authors can consult while writing the manuscript (Keil et al., 2014, section Appendix). Here we propose to anticipate this time investment, with the advantage that carefully thinking about these methodological details before data collection may lead to improvements in the study design when still useful.

Some researchers might also be worried that the time invested in writing the preregistration would be wasted if results do not pan out as expected and, consequently, the final manuscript would be more difficult to publish. Indeed, current incentive structures (be it hiring practices, journals, or funders) usually value quality over quantity (Allen and Mehlert, 2019) and exert pressure to publish novel, groundbreaking, positive results (Fanelli, 2012, 2010; Ioannidis et al., 2014; Jennings and Van Horn, 2012; Nissen et al., 2016; Scargle, 2000). Nonetheless, many academic journals accept manuscripts with non-significant findings if the methodology is robust, with the aim to mitigate the pervasive problem of publication bias.21 Similarly, hiring practices are starting to reward open science practices (Schönbrodt et al., 2020). We are highly sympathetic to Allen and Mehlert’s (2019) call to align incentive structures more with open and transparent research practices that value quality over quantity (see also Flier, 2017).

Unforeseen circumstances – e.g., problems recruiting the planned number of participants due to a pandemic – may require reasonable deviations from the original preregistered plan. This is acceptable as long as it is transparently documented in the published paper. Regrettably, recent evidence shows that undisclosed protocol deviations are common. An analysis of preregistered studies in the journal Psychological Science (Claesen et al., 2019) showed that none of them had perfectly anticipated every step of the research project: differences between preregistration protocols and final manuscripts were observed, for instance, in sample size, exclusion criteria, and statistical models, with only one study transparently reporting all discrepancies. Partial or lack of disclosure of deviations from pre-study plans is a well-known problem not only in social sciences (e.g., Franco et al., 2016) but also in clinical trials (Goldacre et al., 2019), for which registration is compulsory or strongly encouraged in many countries (see the World Health Organisation Registry Network, https://who.int/ictrp/network/en/). We emphatically recommend to clearly report any deviations from the preregistered plan, preferably in a separate section in the main manuscript or in Supplementary Materials. A useful checklist can be found on the OSF (https://osf.io/yrvcg/).

4.3. Preregistration is not a silver bullet

As argued in Section 2.1, preregistration can strengthen the evidential value of studies by increasing transparency, disclosing selective outcome reporting, and increasing the number of publications with non-significant findings. However, adopted in isolation, preregistration is not sufficient to increase scientific rigor; for example, it may not necessarily prompt researchers to carefully examine whether their chosen statistical models are appropriate for the experimental question (Guest and Martin, 2021; Szollosi et al., 2020), improve statistical inferences (Devezer et al., 2020; Navarro, 2020), strengthen the link between theories and their mathematical representations (Szollosi and Donkin, 2019), or develop more precise, consistent, and ‘hard-to-very’ theories altogether (Szollosi and Donkin, 2021; van Rooij and Baggio, 2021). In other words, preregistration in itself does not necessarily improve the quality of the research, and it might even be harmful if it grants statistically invalid or theoretically weak research an unwarranted higher status (“a superficial veneer of rigor”); Devezer et al., 2020, p. 19) compared to non-preregistered, but otherwise solid and transparently documented, research. In fact, when other conditions are fulfilled – e.g., a strong theoretical framework that warrants precise predictions; a clear justification of analytical choices; open data, materials, and code; and/or convergent results via multiverse analysis –, preregistration does not necessarily lead to more robust and trustworthy conclusions (Rubin, 2020). Unfortunately, these conditions are rarely fulfilled in electrophysiology and psychophysiology, even when investigating popular topics with a long research tradition. As an example, a recent review on the electrophysiological correlates of early word prediction (Nieuwland, 2019) analyzed available evidence for the sensory hypothesis as opposed to the recognition hypothesis. This analysis revealed that current published evidence is often obtained via novel tasks with unclear specificity and sensitivity, in samples which might not be sufficient for a precise estimation of small or medium effect sizes, using statistical analyses that do not allow to accurately partition between different sources of variance, and whose data cleaning and analysis procedures are unavailable for scrutiny. As argued in the current paper, preregistration may help researchers think about all of these steps in the planning stage of their projects, with clear advantages in avoiding confirmation and hindsight bias. We thus believe that preregistration serves the important goal of increasing transparency (Navarro, 2020) by offering a window into the research workflow, an often messy and non-linear process that is far from the flawless stories recounted in academic papers. Accepting these imperfections may promote a work culture that normalizes errors, acknowledges the depth of domain-specific knowledge, and fosters intra- and interdisciplinary collaborations (see also Nosek et al., 2012). Moreover, preregistration offers the opportunity to evaluate whether the chosen tests support or falsify theoretical predictions, i.e., their severity (Lakens, 2019).

Finally, just like other research practices, preregistration can in principle be used unethically (Yamada, 2018). For example, one could preregister a large number of similar experiments and keep them under embargo. Whenever one of the studies turns out to be “successful” (i.e., statistically significant), the resulting paper would only refer to the corresponding preregistration and all the other “unsuccessful” ones could be withdrawn. However, metadata and a justification for each withdrawal would still be publicly available and therefore raise suspicion. Another unethical practice has been termed “preregistering after the results are known” (PARKing; Yamada, 2018), i.e., drafting and publishing the preregistration of a study that has already been completed and whose results conveniently fit within the narrative of the “preregistered” document. Complementary open science practices, e.g., data and code sharing, can effectively mitigate this risk. Journal editors and reviewers are invited to carefully compare the preregistered document with the manuscript during submission and evaluation, as well as request raw data and analysis code by default (Morey et al., 2016). Having said that, we prefer to think of fellow researchers in more

20 https://antonio-schettino.com/post/2019-07-23-prereg-challenge/.

21 See, for instance, two recent Nature (2020, 2017) editorials, the PLOS ONE article collection Missing Pieces (2015), the editorial by Munafò and Neill (2016) in the Journal of Psychopharmacology, as well as the submission guidelines of Meta-Psychology (https://open.lmu.de/index.php/metapsychology/about) and Royal Society Open Science (https://royalsocietypublishing.org/rrss/s-for-authors).
optimistic terms, motivated by higher goals than simply publishing as much as possible (although we might, admittedly, be hopelessly naive; Chapman et al., 2019; DeDeo, 2020). We also point out that premediated approaches to exploit the vulnerabilities of a system can hardly be reconciled with claims that mistakes were made in good faith due to ignorance or procedural complexity: researchers engaging in such behaviors (if proven) would consciously commit fraud, and responsible institutions should be contacted and deliver appropriate sanctions.

5. Conclusion

The adoption of a new procedure can be met with resistance, particularly if the benefits are unclear, the amount of work is perceived as too onerous, and training and guidance are lacking. In this paper, we argue that preregistering EEG projects can effectively facilitate the transparent reporting of data preprocessing and analysis choices, thereby improving study replicability and the verifiability of published knowledge. The time spent writing a preregistration is saved at later stages, because the information included in a comprehensive protocol is required at the time of publication. Ready-made templates can serve as useful guidelines and facilitate the implementation of this practice in the research workflow. Combined with other open science practices—e.g., sharing study protocols, materials, raw data, and analysis code—preregistration increases transparency in the research process and trustworthiness of the scholarly products not only for academic peers, but also other stakeholders in society (Jamieson et al., 2019).

Contributions

AS supervised the project. All authors wrote the initial draft, reviewed, edited, and approved the final version of the manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

We would like to thank Dr. Matt Craddock for his input on a preliminary draft, as well as the kind and thoughtful comments by Dr. Mante S. Nieuwen and another anonymous reviewer on an earlier version of the manuscript.

Declaration of competing interest

MP is funded by the Max Planck Society and the Berlin School of Mind and Brain. GHG is funded by the Max Planck Society and the Einstein Center for Neurosciences Berlin. AS is employed at Erasmus University Rotterdam as Coordinator Open Science. Funders and employers had no role in project design, literature writing, review of the paper, and/or decision to submit for publication.

References

Adda, J., Decker, C., Ottaviani, M., 2020. P-hacking in clinical trials and how incentives paper, and/or decision to submit for publication.

M. Paul et al. International Journal of Psychophysiology 164 (2021) 52–63
Rewarding negative results keeps science on track. Nature 551, 2017, 414. https://doi.org/10.1038/nature24656.

van Rooij, I., Baggio, G., 2021. Theory Before the Test: How to Build High-Verisimilitude Explanatory Theories in Psychological Science. Perspect. Psychol. Sci. https://doi.org/10.1177/1745691620970604.

Rosenthal, R., 1979. The file drawer problem and tolerance for null results. Psychol. Bull. 86, 628–641. https://doi.org/10.1037/0033-2909.86.6.628.

Rouder, J.N., 2014. Optional stopping: no problem for Bayesians. Psychon. Bull. Rev. 21, 301–308 (doi:10/fssxap).

Rouder, J.N., Speckman, P.L., Sun, D., Morey, R.D., Verson, G., 2009. Bayesian t tests for accepting and rejecting the null hypothesis. Psychon. Bull. Rev. 16, 225–237 (doi:10/bshd3p).

Rousselet, G.A., 2012. Does filtering preclude us from studying ERP time-courses? Front. Psychol. 3 (doi:10/ggwb6g).

Rubin, M., 2020. Does preregistration improve the credibility of research findings? Quant. Methods Psychol. 16, 376–390 (doi:10/gzw29j).

Sander, A., Banica, I., Riesel, A., Flake, J., Klawohn, J., Weinberg, A., 2020. Comparing the effects of different methodological decisions on the error-related negativity and its association with behaviour and gender. Int. J. Psychophysiol. 156, 18–39. https://doi.org/10.1016/j.ijpsycho.2020.06.016.

Sarge, J., 2000. Publication bias: the ‘File-Drawer’ problem in scientific inference. J. Sci. Explor. 14, 91–105.

Scheel, A.M., Schijen, M., Lakens, D., 2020. An Excess of Positive Results: Comparing the Existing Literature (Preprint). PsyArXiv. https://doi.org/10.31234/osf.io/jp6wy.

Van der Akker, O., Weston, S.J., Campbell, L., Chapik, W.J., Damian, R.L., Davis-Kean, P., Hall, A.N., Koscie, J.E., Krue, E.T., Olsen, J., Ritchie, S.J., Valentine, K.D., van’t Veer, A.E., Bakker, M., 2019. Preregistration of Secondary Data Analysis: A Template and Tutorial (Preprint). PsyArXiv. https://doi.org/10.31234/osf.io/brfrm.

Van Doors, J., van den Bergh, D., Böhm, U., et al., 2020. The JASP guidelines for conducting and reporting a Bayesian analysis. Psychol. Bull. Rev. https://doi.org/10.3758/s13423-020-01798-5.

van Murn, R., Ute, S.J., Dall’Aglio, G., Blok, E., Schettino, A., Schramm, L., Weber, B., Tangermann, M., 2011. Automatic classification of artifactual ICA-results: implications of baseline potentials and variability with application to amplitude normalization by vector scaling. Biol. Psychol. 72, 333–343. https://doi.org/10.1016/j.biopsycho.2005.11.012.

Van den Akker, O., Weston, S.J., Campbell, L., Chapik, W.J., Damian, R.L., Davis-Kean, P., Hall, A.N., Koscie, J.E., Krue, E.T., Olsen, J., Ritchie, S.J., Valentine, K.D., van’t Veer, A.E., Bakker, M., 2019. Preregistration of Secondary Data Analysis: A Template and Tutorial (Preprint). PsyArXiv. https://doi.org/10.31234/osf.io/brfrm.

Wagenmakers, E.J., 2007. A practical solution to the pervasive problems of p values. Psychol. Bull. Rev. 14, 779–804 (doi:10/gww2q).