

UC Davis

UC Davis Previously Published Works

Title

Good scientific practice in EEG and MEG research: Progress and perspectives

Permalink

<https://escholarship.org/uc/item/2866940q>

Authors

Niso, Guiomar

Krol, Laurens R

Combrisson, Etienne

et al.

Publication Date

2022-08-01

DOI

10.1016/j.neuroimage.2022.119056

Peer reviewed



Published in final edited form as:

Neuroimage. 2022 August 15; 257: 119056. doi:10.1016/j.neuroimage.2022.119056.

Good scientific practice in EEG and MEG research: Progress and perspectives

Guiomar Niso^{a,b,1}, Laurens R Krol^{c,1}, Etienne Combrisson^d, A. Sophie Dubarry^e, Madison A Elliott^f, Clément François^e, Ysult Héjja-Brichard^g, Sophie K Herbst^h, Karim Jerbi^{i,j}, Vanja Kovic^k, Katia Lehongre^l, Steven J. Luck^m, Manuel Mercierⁿ, John C. Mosher^o, Yuri G Pavlov^{p,q}, Aina Puce^a, Antonio Schettino^{r,s}, Daniele Schönⁿ, Walter Sinnott-Armstrong^t, Bertille Somon^u, An ela Šoški^{k,v}, Suzy J. Styles^{w,x}, Roni Tibon^{y,z}, Martina G. Vilas^{aa}, Marijn van Vliet^{ab}, Maximilien Chaumon^{l,*}

^aPsychological & Brain Sciences, Indiana University, Bloomington, IN, USA

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

*Corresponding author. maximilien.chaumon@icm-institute.org (M. Chaumon).

¹These authors contributed equally to this work.

Declaration of Competing Interest

The authors declare no conflict of interest related to this work.

Citation diversity statement

Recent work in several fields of science has identified a bias in citation practices such that papers from women and other minority scholars are under-cited relative to the number of such papers in the field (Bertolero et al., 2020; Chatterjee and Werner, 2021; Dworkin et al., 2020; Fulvio et al., 2021). Here we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, race, ethnicity, and other factors. First, we obtained the predicted gender of the first and last author of each reference by using databases that store the probability of a first name being carried by a woman (Dworkin et al., 2020; Zhou et al., 2020). Expected proportions estimated from five top neuroscience journals since 1997 are 6.7% woman(first)/woman(last), 9.4% man/woman, 25.5% woman/man, and 58.4% man/man (Dworkin et al., 2020). By this measure, our references contain 10.42% woman/woman, 13.57% man/woman, 15.61% woman/man, and 60.4% man/man. This method is limited in that (a) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity and (b) it cannot account for intersex, non-binary, or transgender people. Second, we obtained the predicted racial/ethnic category of the first and last author of each reference by databases that store the probability of a first and last name being carried by an author of color (Ambekar et al., 2009; Sood and Laohaprapanon, 2018). Using a similar random draw model regressing for relevant variables, the expected race proportions in reference lists were 11.9% for author-of-color/author-of-color, 12.8% for white/author-of-color, 23.5% for author-of-color/white, and 51.8% for white/white (Bertolero et al., 2020). By this measure, our references contain 7.23% author-of-color/author-of-color, 15.48% white/author-of-color, 17.31% author-of-color/white, and 59.98% white/white. This method is limited in that (a) names and Florida Voter Data to make the predictions may not be indicative of racial/ethnic identity, and (b) it cannot account for Indigenous and mixed-race authors, or those who may face differential biases due to the ambiguous racialization or ethnicization of their names. We look forward to future work that could help us to better understand how to support equitable practices in science.

Credit authorship contribution statement

Guiomar Niso: Conceptualization, Project administration, Supervision, Writing – original draft, Writing – review & editing. **Laurens R. Krol:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing. **Etienne Combrisson:** Writing – original draft, Writing – review & editing. **A. Sophie Dubarry:** Conceptualization, Writing – original draft, Writing – review & editing. **Madison A. Elliott:** Writing – original draft, Writing – review & editing. **Clément François:** Conceptualization. **Ysult Héjja-Brichard:** Writing – original draft, Writing – review & editing. **Sophie K. Herbst:** Writing – original draft, Writing – review & editing. **Karim Jerbi:** Writing – original draft, Writing – review & editing. **Vanja Kovic:** Writing – original draft, Writing – review & editing. **Katia Lehongre:** Writing – original draft, Writing – review & editing. **Steven J. Luck:** Writing – original draft, Writing – review & editing. **Manuel Mercier:** Conceptualization, Writing – review & editing. **John C. Mosher:** Writing – original draft, Writing – review & editing. **Yuri G. Pavlov:** Writing – original draft, Writing – review & editing. **Aina Puce:** Writing – original draft, Writing – review & editing. **Antonio Schettino:** Writing – original draft, Writing – review & editing. **Daniele Schön:** Writing – original draft, Writing – review & editing. **Walter Sinnott-Armstrong:** Writing – original draft, Writing – review & editing. **Bertille Somon:** Writing – original draft, Writing – review & editing. **An ela Šoški:** Writing – original draft, Writing – review & editing. **Suzy J. Styles:** Writing – original draft, Writing – review & editing. **Roni Tibon:** Writing – original draft, Writing – review & editing. **Martina G. Vilas:** Writing – original draft, Writing – review & editing. **Marijn van Vliet:** Writing – original draft, Writing – review & editing. **Maximilien Chaumon:** Conceptualization, Project administration, Supervision, Writing – original draft, Writing – review & editing.

^bUniversidad Politecnica de Madrid and CIBER-BBN, Madrid, Spain

^cNeuroadaptive Human-Computer Interaction, Brandenburg University of Technology Cottbus-Senftenberg, Germany

^dAix-Marseille University, Institut de Neurosciences de la Timone, France

^eCNRS, LPL, Aix-Marseille University, Aix-en-Provence, France

^fThe University of British Columbia, Canada

^gCentre d'Ecologie Fonctionnelle et Evolutive, CNRS, EPHE, IRD, Université Montpellier, Montpellier, France

^hCognitive Neuroimaging Unit, INSERM, CEA, CNRS, NeuroSpin center, Université Paris-Saclay, Gif/Yvette, France

ⁱCognitive and Computational Neuroscience Laboratory, Department of Psychology, University of Montreal, Montreal, QC, Canada

^jMila - Quebec Artificial Intelligence Institute, Canada

^kFaculty of Philosophy, Laboratory for neurocognition and applied cognition, University of Belgrade, Serbia

^lInstitut du Cerveau - Paris Brain Institute - ICM, Inserm U 1127, CNRS UMR 7225, APHP, Hôpital de la Pitié Salpêtrière, Sorbonne Université, Centre MEG-EEG, Centre de NeuroImagerie Recherche (CENIR), Paris, France.

^mCenter for Mind & Brain, University of California, Davis, CA, USA

ⁿAix Marseille Univ, Inserm, INS, Inst Neurosci Syst, Marseille, France

^oMcGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX, USA

^pUniversity of Tuebingen, Germany

^qUral Federal University, Yekaterinburg, Russia

^rErasmus University Rotterdam, Rotterdam, the Netherland

^sInstitute for Globally Distributed Open Research and Education (IGDORE), Sweden

^tDuke University, USA

^uISAE-SUPAERO, Université de Toulouse, France

^vTeacher Education Faculty, University of Belgrade, Serbia

^wPsychology, Nanyang Technological University, Singapore

^xSingapore Institute for Clinical Sciences, A * STAR, Singapore

^yMRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK

^zSchool of Psychology, University of Nottingham, Nottingham, UK

^{aa} Ernst Strüngmann Institute for Neuroscience, Frankfurt am Main, Germany

^{ab} Aalto University, Finland

Abstract

Good scientific practice (GSP) refers to both explicit and implicit rules, recommendations, and guidelines that help scientists to produce work that is of the highest quality at any given time, and to efficiently share that work with the community for further scrutiny or utilization. For experimental research using magneto- and electroencephalography (MEEG), GSP includes specific standards and guidelines for technical competence, which are periodically updated and adapted to new findings. However, GSP also needs to be regularly revisited in a broader light. At the LiveMEEG 2020 conference, a reflection on GSP was fostered that included explicitly documented guidelines and technical advances, but also emphasized intangible GSP: a general awareness of personal, organizational, and societal realities and how they can influence MEEG research. This article provides an extensive report on most of the LiveMEEG contributions and new literature, with the additional aim to synthesize ongoing cultural changes in GSP. It first covers GSP with respect to cognitive biases and logical fallacies, pre-registration as a tool to avoid those and other early pitfalls, and a number of resources to enable collaborative and reproducible research as a general approach to minimize misconceptions. Second, it covers GSP with respect to data acquisition, analysis, reporting, and sharing, including new tools and frameworks to support collaborative work. Finally, GSP is considered in light of ethical implications of MEEG research and the resulting responsibility that scientists have to engage with societal challenges. Considering among other things the benefits of peer review and open access at all stages, the need to coordinate larger international projects, the complexity of MEEG subject matter, and today's prioritization of fairness, privacy, and the environment, we find that current GSP tends to favor collective and cooperative work, for both scientific and for societal reasons.

Keywords

Magnetoencephalography (MEG); Electroencephalography (EEG); Good scientific practice

1. Introduction

The generation of scientific knowledge, which is the purview of science and the humanities, relies on scientists and researchers using appropriate tools and techniques, appropriate processes and methods, and appropriate ways of thinking and reasoning. All these elements of scientific work have been developed and refined over centuries of discovery to ensure that the generated knowledge lives up to ever-evolving current standards. Rules, guidelines, and principles that codify this way of work are referred to as good scientific practice (GSP). The details vary between disciplines, but generally GSP serves to uphold and continue methodological progress, to make sure results are reliable, and to avoid scientific misconduct.

In magnetoencephalography (MEG) and electroencephalography (EEG), here collectively referred to as MEEG, the first guidelines for good practice were published for EEG by Donchin et al. (1977). They represented the findings of an international committee formed to provide publication criteria, which the committee members first discussed among

themselves and then opened up for public discussion during a conference. It was here that the recommendation was recorded to, among other things, report age, sex, and handedness of all participants – something still done to this day. Since then, this has been a common pattern: new or additional guidelines have often emerged from international meetings or societies where scientists consider, for example, the need for standardization in light of new methods, techniques, or findings.

Currently, the most recent effort in the MEEG community to coordinate and promote GSP at all stages of the research cycle has come from the Organization for Human Brain Mapping. Their COBIDAS MEEG white paper (Pernet et al., 2020) lists current standards and good practices for data acquisition, analysis, reporting, and sharing. Together with numerous other guidelines that have emerged over the years (Donchin et al., 1977; Pivik et al., 1993; Picton et al., 2000; Handy, 2005; Luck, 2005, 2014; Duncan et al., 2009; Gross et al., 2013; Keil et al., 2014; Kappenman and Luck, 2016; Hari and Puce, 2017; Kane et al., 2017; Hari et al., 2018), they support scientists and researchers in minimizing known pitfalls and adhering to best practices during the various stages of research.

Such documents are even more important now that the field is becoming increasingly interdisciplinary, and not all colleagues will have had the opportunity to be trained in all relevant technologies (Pernet et al., 2020). Furthermore, recent concerns about reliability in neuroimaging research (Button et al., 2013; Poldrack et al., 2017), more pressingly referred to as a “replication crisis” (Shrout and Rodgers, 2018), have highlighted the importance of community standards – not just their existence, but also the community’s own awareness of, and adherence to, such standards. Therefore, the field of neuroimaging in general is actively and effectively working towards consolidating GSP within its community (Poldrack et al., 2020).

The virtual LiveMEEG conference on “Good scientific practices in EEG and MEG Research” was held on October 5–9, 2020, with the aim of bringing together MEEG experts to discuss essential aspects of GSP for the entire lifecycle of MEEG research projects. Importantly, what emerged from this meeting was that not everything that is deemed important is captured by current or previous standards – nor, indeed, could it be: there are many subtleties and approaches that do not lend themselves to being standardized per se. Instead, this different type of GSP enables a more general reflection on the way we think and act, with respect to both the work itself and the broader scientific environment. In particular, “open science” is seen as one fundamental aspect of modern GSP. This refers to a general approach which aims to make all products of scientific work publicly available. Among other things, this makes science less isolated, more community-driven, and allows the results to be fully reproduced and verified by colleagues (Garrett-Ruffin et al., 2021; Clayson et al., 2022). As we shall see, this approach is gaining traction, and many modern recommendations follow almost naturally from this general, open mindset.

Our two-fold aim here is to review existing and developing guidelines and resources for GSP, and to capture those LiveMEEG contributions that widened our understanding of what GSP means. GSP is not merely about adhering to established protocols and avoiding mistakes: it involves a more general awareness of personal, organizational, and societal

realities, of the structures that surround and influence us all, and of the future that we ourselves wish to see on the path of scientific progress. This work aims to help foster that awareness.

1.1. How to use this paper

Part of the nature of general principles is that they can be formulated broadly. We do this at the beginning of each section, and use the subsections and subsubsections to provide increasingly detailed GSP that illustrate these principles. To that end, each section first summarizes the general considerations addressed in that section, and briefly introduces its subsections. These subsections deal with more detailed but still generally relevant GSP, while any deeper sections address concrete issues and provide specific suggestions. As such, the reader can use the higher levels to navigate the paper, and choose to read those lower levels as they see fit. The three main sections of this paper, following this introduction, relate to: GSP relevant to the early, e.g., planning and pre-planning stages of MEEG research (Section 2); issues that arise and GSP that can help during data collection, analysis, and reporting (Section 3); and finally, considerations and GSP that go beyond the work itself, and touch upon ethical and social aspects that surround it (Section 4). The reader is encouraged to look into the referenced literature in any relevant section, as well as Appendix Table A1 for further resources.

2. Early considerations for MEEG projects

Some general principles of GSP come into play even before the first instrument is picked up. These include GSP specifically related to planning and pre-planning stages of a project, as well as GSP that is not specific to any one stage, but rather reflects a general approach to the project as a whole.

During pre-planning stages, among other things, one may be considering the outcomes of previous experiments and how they may be followed up on, or brainstorming about experimental variables and manipulations. Here, it is important to be aware of the limits of one's own ability to properly seek out, perceive, and arrange the relevant facts. Section 2.1 deals with GSP related to this. It emphasizes the importance of awareness, and of availing oneself of the expertise of others early on. This latter point also applies to planning stages, where there are a number of benefits of sharing project ideas with others, both before and after they have been finalized. To that end, Section 2.2 explains the various facets of public pre-registration. Finally, open science was mentioned in the introduction as a fundamental approach to the organization and conducting of scientific projects. Open science requires a commitment that impacts all stages of research, but many tools, guides, and examples already exist to support this, some of which are outlined in Section 2.3.

2.1. Human factors

Scientists are trained to form beliefs about the world and test their theoretical and experimental hypotheses with objectivity. However, even expert humans are fallible, as they can be victims of their own biases and, consequently, make logical errors. GSP can help mitigate these errors by making researchers aware of them, by explicitly accounting for

biases in data analysis and representation, and by enforcing careful planning of experiments, with attention to the logical implications of given observations.

2.1.1. Cognitive biases, their manifestation, and ways to mitigate them—

Cognitive biases pervade all levels of thinking and reasoning, from low level perceptual processes to social representations. Because cognitive capacities are limited, humans allocate resources sparingly and create mental shortcuts (heuristics) that are prone to oversights and biases (Gigerenzer, 2008). Of note here are the biases that cause researchers to perceive or understand information incorrectly during data analysis and viewing, and end up drawing inaccurate conclusions. We illustrate a few effects and mention strategies to help overcome these biases.

At the perceptual level, researchers have to grapple with sensory errors and visual illusions both when reading and when creating visual data representations. The limits and strengths of the visual system for extracting information from such visual representations are well studied, and detailed guidelines exist: see Franconeri et al. (2021) for a recent review.

A prominent example in the field of brain imaging in general, and MEEG in particular, is the question of color mapping. Because human perception of color categories is neither linear nor equal across the color wheel (Bae et al., 2014), the choice of how to represent continuous data with color is not trivial. The rainbow color map, widely used in the field, unfortunately does not take perceptual limitations into account and can lead to perceiving inaccurate steps in the data, or meaningless categorical boundaries (Cooper et al., this issue; Borland and Taylor II, 2007). Instead, there are now perceptually uniform colormaps that ensure that the linear representation of data is correctly perceived by human observers. See Cooper et al. (this issue), for a review and suggestions. Another example is how results can be distorted when not visualized properly, hiding underlying patterns in the raw data (Chambers et al., 2018; Ware, 2019; Allen et al., 2021). Therefore, it is recommended to present more complete visual representations such as box and density plots that reflect the entire distribution of the underlying data rather than focusing on its central tendency (Hintze and Nelson, 1998; Rodu and Kafadar, 2021).

Beyond individual perceptual errors, other types of errors can occur when viewing data in a social context. A well-known example is the “curse of knowledge” effect, whereby personal experience and familiarity with a method, theory, or set of results make the presenter assume more knowledge in the audience than there actually is Xiong et al. (2020). This is almost always the case when researchers present their own work, and may thus lead to a miscommunication between presenter and observer. This disparity may be further increased due to the fact that, on the observer’s side, their high-level experiences influence their interpretation and even low-level perceptual processing of objective information. This can affect the reading and understanding of MEEG data, especially when it is likely that an experimental result strongly supports or refutes one’s own theoretical predictions. In combination with the confirmation bias, the same viewer may be inclined to ignore data features that refute their own theoretical views. There is a plethora of higher-order cognitive biases that affect decision making at many levels. Commenting on all of them is beyond the scope of the present article, but their powerful effect on perception and understanding is

important to consider at all stages of a research project. The interested reader could consult references such as Baron (2006) or the seminal works of Tversky and Kahneman (1974).

Finally, social factors also influence the socio-cultural structure of scientific investigation, and influence scientific progress in a way that may be biased toward theories and representations of the most-represented social groups. The majority of publications in MEEG research come from Western institutions that recruit so-called W.E.I.R.D. (Western, Educated, Industrialized, Rich, and Democratic) participants, putting into question whether these study results generalize to a larger population (Henrich et al., 2010). Researchers must keep this in mind when reading past work, designing new studies, and describing their own experimental conclusions. Lack of diversity and inclusion is prevalent in academia, and the MEEG field is no exception to this (Schrouff et al., 2019). Gender bias, as well as marginalization and underrepresentation of racial, ethnic, and cultural minorities, affect women, BIPOC (black, indigenous, and other people of color) and people with disabilities, not only as students and scientists, but also in participant samples. Some relevant resources and concrete actions to mitigate gender bias and inequity at individual and institutional levels in academia have been collected recently (Schreiweis et al., 2019; Llorens et al., 2021; Levitis et al., 2021). As suggested by Dworkin et al. (2020), Zurn et al. (2020) we have included a citation diversity statement at the end of the manuscript for the awareness of citation bias in the literature. In the field of EEG, recent studies have begun to address the fact that standard EEG electrodes are not designed to accommodate coarse and curly hair common in individuals of African descent (Etienne et al., 2020; Choy et al., 2021). This is an important reminder that systemic bias can also occur through instrumentation.

2.1.2. Logical thinking and fallacious reasoning: an example—Cognitive biases affect perception, cognition, and reasoning in science, but fallacious conclusions can also be reached due to a failure to fully understand and account for the logical implications of an experimental work. Logical thinking is at the heart of conceiving hypotheses and designing experiments. Experimenters need to take certain steps at the outset to forestall later mistakes when drawing conclusions from results. See Sinnott-Armstrong and Simmons (2021) for a list of common fallacies in MEEG data.

Some common mistakes can be illustrated by misinterpretations of Libet and collaborators (Libet et al., 1983). In this now infamous EEG experiment, participants were positioned in front of a clock face, and were instructed to freely move their fingers or wrists at a time of their choosing. When they did decide to move, they were asked to report the time when the decision was made. As such, the experiment yielded three points in time: the self-reported time of the conscious will or decision to move (W), the measured time that movement indeed took place (M), and correlated brain activity in the EEG in the form of a readiness potential (RP). They found that the readiness potential RP occurred *before* the reported time of W, which in turn occurred before M. Libet's findings have been replicated often, but there are still raging debates about their interpretation, the role of the neural activity (RP) and W in producing M, and any relation of these findings to free will. Unfortunately, these debates continue to involve several fallacies that could have been avoided by following these simple recommendations.

First, spell out all theories or hypotheses that might explain possible experimental findings. Overlooking an option can lead to a fallacy called *false dichotomy*. For example, commentators on Libet et al. (1983) often assumed that one of only three options must be true: (a) W causes M, but nothing causes W (Libertarianism); (b) W causes an RP, which causes M (Backwards Causation); or (c) RP causes M, but W does not cause M (Libet's conclusion). Libet's followers rejected the Libertarian claim, because they assumed that every event has a cause. They also rejected Backwards Causation, because W occurs later than RP, and causes never come after their effects. They concluded that RP causes M, but W does not cause M. What they failed to consider was the Commonsense view that RP causes W, and then W in turn causes M—that is, activity in the brain causes choices, which cause actions. To overlook such a plausible option is called *false dichotomy* because it is false to assume that one member of the original set of alternatives must be true.

Second, think carefully about which theories or hypotheses are logically compatible with others. Forgetting that both of two supposed alternatives might be true can lead to a fallacy called *affirming a disjunct*. Some readers of Libet et al. (1983) seem to reason that either RP causes M or W causes M, so, if RP causes M, then W does not cause M (cf Sinnott-Armstrong, 2011). However, this reasoning is fallacious because RP and W might both cause M, such as when RP, W, and M form a chain of causes across time.

Third, define the precise category that the conclusion will cover and include stimuli or tasks for all variations within that category. Failing to include a subclass of the phenomenon can lead to a fallacy called *hasty generalization*. An experiment using a specific kind of action cannot justify a conclusion about all actions, especially when the studied acts are atypical. In particular, Libet's subjects arbitrarily chose when to flex their wrists with nothing at stake. What Libet found about these arbitrary, simple, trivial acts might not hold for important human actions that are based on conscious deliberation and require complex sequences. If one wants to draw a conclusion about all actions, one must be sure to include diverse kinds of actions in the experimental data itself.

How can one know whether one has included enough kinds of test alternatives, considered enough theories, and specified competitors that cannot both be true? The best way to minimize the chances of these and other mistakes is to seek early feedback from others, especially those who disagree with one's views and assumptions. This includes people from other fields and backgrounds. Working together instead of alone is of great benefit: today's team science activities have shown these advantages (see examples in Section 2.3).

2.1.3. Example strategy to minimize biases and errors—Experiencing cognitive biases or logical fallacies and having them influence research at least in part may be inevitable, but it is possible to diminish their impact on research practices by learning about them and by using concerted “slow” critical thinking skills (Kahneman, 2013). Appendix Table A1 has a few current references on ways to characterize and circumvent biases and errors.

One possible strategy is to organize a “premortem meeting” with fellow researchers before starting a new project (Klein, 2007). This exercise involves simulating a future

meeting, assuming the project has failed, and then working backwards from that outcome to determine what could have caused that “failure”. This meeting can include a focused conversation about existing biases and beliefs within the group that might influence the work, and then use imaginative exercises to collaboratively solve these issues. As researchers plan and implement their study, it is important to keep track of which methods are being chosen and why (see also Section 2.2. below). At this stage, researchers are encouraged to document which rationales may be biased and how. Skepticism is healthy in science; listening to data is critical, and trying to understand both how data do, or do not fit in with existing theories is GSP. A useful set of guidelines come from Abelson’s *MAGIC Criteria* (Abelson, 1995), which can be used to understand the *Magnitude, Articulation, Generality, Interestingness, and Credibility* of statistical claims in research. For the many reasons listed in the chapter, once researchers are comfortable operating within a scientific paradigm, it becomes especially challenging to shift their beliefs. But in order to progress science fairly, they must be willing to do so.

2.2. Pre-registration

The primary purpose of GSP during the planning stage of a scientific project is to ensure the correct outcome of that project, i.e., that valid conclusions can be drawn from the collected data. All phases of the project must be considered and aligned for this to be possible. Beyond the aspects mentioned above in Section 2.1, the many factors involved in MEEG projects make for a complex web of interdependent decisions that need to be taken. Pre-registration, i.e. the practice of defining and (eventually) publicly disclosing experimental plans before data are collected and/or analyzed, can help in this process in at least two ways that we describe below, before highlighting different pre-registration procedures and tackling possible concerns.

First, MEEG analysis pipelines are particularly complex, combining a multitude of pre-processing and analytical steps, each of which involves numerous parameters. Together, these result in a large number of combinations to choose from. This problem is known in the literature as the “garden of forking paths” (Gelman and Loken, 2013). Given that it is difficult to test the independent contribution of each combination (see Section 3.4.4), novice and even expert neuroscientists are often left with not knowing how strongly a given effect (or the absence thereof) depends on their analysis choices. Without detailed planning in advance, researchers often make data-dependent choices, which undermines to some extent the logic of statistical tests designed for a priori set hypotheses (Gelman and Loken, 2013; Luck and Gaspelin, 2017). Pre-registration helps researchers to carefully plan the analysis “path” that best fits the goals of the experiment.

Second, sometimes, the only valid conclusion is a null result. Unexpected, ambiguous, or null results are an essential part of scientific discovery. However, the academic incentive system rewards conclusive results and concise stories with publications in “high-impact” journals. “Negative” results are often not publishable, resulting in the infamous “file drawer” problem (Rosenthal, 1979) and the effort of the researcher remaining unrecognized. This mismatch between the academic incentive system and the reality of scientific work (“getting it published” rather than “getting it right”; Nosek et al., 2012) leads to publication bias

and unreliable practices in published work. This is reflected, for example, in a peculiar prevalence of p -values just below the critical (yet arbitrary) cut-off of 0.05 (Masicampo and Lalande, 2012) and inflated effect sizes in published studies (Button et al., 2013). A move towards increased transparency can help mitigate publication bias, selective reporting, and suboptimal research practices, and reward accurate work instead of striking results. Pre-registration is one option to make this happen. Initial evidence suggests that the research quality of pre-registered studies is judged higher than that of traditional publications (Soderberg et al., 2021), while a comparison of registered versus conventional reports revealed a striking imbalance of 44% versus 96% of positive findings, respectively (Scheel et al., 2021).

2.2.1. Different types of pre-registration—As mentioned above, pre-registration is the practice of defining and disclosing experimental plans before data are collected or analyzed. Such (public) disclosure of plans allows for them to be scrutinized and updated before they are executed, essentially inviting peer review not just at the final stages of the research project (i.e., publication) but throughout the entire lifecycle of the scientific work. This is facilitated by online platforms that allow pre-registration of work during preliminary stages and allow revision of the initial protocol (see Appendix Table A1). Another recently suggested format for this purpose is “prereg posters”, i.e., conference posters that present planned scientific projects (Tibon et al., 2018; Brouwers et al., 2020). We use the term “**unlocked pre-registration**” to refer to this process of depositing planned work in the public domain and calling for feedback from fellow researchers. At this preliminary stage, the project’s experimental design and analytical approaches are openly discussed, reviewed, and adjusted in the general scientific arenas. These aspects of the project remain “unlocked” and are subject to changes following peer feedback. Regardless of the platform used, presenting research plans prior to data collection allows researchers to receive feedback on their hypotheses, design, and analyses from colleagues, a process that will likely improve the study.

Once the protocol is finalized, it is amenable to a formal pre-registration, i.e. a time-stamped protocol on a public website (see Appendix Table A1) which describes the experimental and analytical procedures of the study in detail. We use the term “**locked pre-registration**” to indicate that, at this point, the protocol is finalized in principle and that any deviation will have to be clearly justified and documented in the final research paper. This type of pre-registration can help emphasize the planning phase of a study (also as an educational resource for trainees), increase procedural transparency, and serve to publicly take credit for an ongoing study, hypothesis, or theoretical model even before the results are in.

A study protocol can also be locked via the submission of a “**Registered Report**”. This is a published article format that was pioneered by the journal *Cortex* in 2013 (Chambers, 2013) and has since been implemented by more than 250 journals (Nosek et al., 2018; Chambers and Tzavella, 2022). In contrast to the pre-registration options discussed above, study protocols submitted as Registered Reports systematically undergo editorial triage and peer-review before data collection starts. In the case of “in principle acceptance” (stage 1), the journal commits to publishing the final report irrespective of whether the results match the initial hypotheses. In a second phase, following data collection and analysis (stage

2), the full study is submitted and reviewed again, with guaranteed publication as long as the approved protocol was followed (with changes sufficiently documented and justified). Registered Reports are beneficial for individual researchers as they motivate and facilitate thorough and careful planning, and provide the opportunity to receive expert feedback on the research plan before data collection. They further ensure acceptance for publication, independent of statistically significant results. For the scientific community in general, they increase the transparency of the scientific process, reduce suboptimal scientific practices (although the impact of pre-registration alone on such practices is debated (Devezer et al., 2020; Rubin, 2020)), and reduce publication bias in the scientific literature.

2.2.2. Potential concerns regarding pre-registration—Despite the benefits of the pre-registration approaches outlined above, there are some commonly raised concerns. First, “publicly pre-registered protocols might be scooped”. However, with locked pre-registrations and Registered Reports, the time-stamped official report provides proof of when the study was originated by the researcher, and should therefore alleviate these concerns. In this respect, pre-registration could be particularly beneficial for research teams working in competitive fields or with fewer resources.

Second, “the project’s progress incurs a delay”. Namely, for Registered Reports, the review process for stage 1 might add months before data collection can start, which might not be feasible for short project durations or for projects investigating an acute phenomenon. In this case, other forms of pre-registration discussed above (posters, locked pre-registration on a dedicated website) might be more appropriate. However, when time permits, this can also be turned into a benefit, as it shifts some of the heavy load from the publication phase to the planning phase.

Third, “proficiency is required for all types of pre-registration”. Coming up with a detailed analysis protocol, especially for complex methodologies like MEEG, requires considerable experience in making informed decisions and utilizing existing pipelines (perhaps from other laboratories; Paul et al., 2021). Nevertheless, pre-registration can still benefit trainees and inexperienced researchers, with support from their mentors and supervisors. For instance, they can gain the necessary expertise by starting with a replication of a previously published study, where experimental and analysis parameters are more easily derived. Moreover, the complete preprocessing pipeline (and some of the analyses) can be performed on pilot data, obtained prior to pre-registration. Standardized study pipelines (see Table 1), templates (e.g., Section 3.6.2), and guidelines (e.g., Paul et al., 2021) are also excellent means to narrow down the possibilities in MEEG analysis (a few more examples are also discussed in Section 3.6). Thus, a research lifecycle that includes pre-registration provides methodological and research training opportunities, even prior to the pre-registration step.

Finally, some might worry that “such a strict format would hinder the creative process of science”. It is important to note that locked pre-registrations and Registered Reports *do not prevent exploratory analyses or reporting of unexpected findings*. The only requirement is that any addition or change to the original protocol is adequately labeled as such, so as to keep track of the original idea and analysis plan in the published report.

Despite their merits, pre-registration and Registered Reports are not the only answer to the replication crisis, nor do they guarantee scientific integrity. However, together with other good practices such as standardized pipelines and open data and code, they are an important tool for improving the transparency and credibility of published science.

2.3. Reproducibility and collaboration: some examples and resources

GSP surrounding open science and reproducibility aims to produce results that can be verified by others (Niso et al., 2022). “Reproducible research” here refers to different but related concepts (Box 2), which all increase the reliability of the observations and conclusions, in the sense that they are independent of who makes them. This goes hand in hand with open practices, which make the processes and materials public. Numerous resources exist to aid researchers in implementing open science and to conduct reproducible research, some of which are listed below in this section.

Working reproducibly and openly is not just a matter of deontology, or higher morals. There are also “selfish reasons” to pursue reproducibility (Markowetz, 2015). For one, it promotes careful bookkeeping so that not just others, but also researchers’ “future self” will be able to understand the work done in as much detail as possible, as easily as possible. Ensuring open access to pipelines as early as possible is also a good way of obtaining feedback from the community, and therefore detect potential mistakes or errors and fix them earlier. Furthermore, working openly and reproducibly helps write papers more efficiently, especially results sections (Markowetz, 2015), provides a citation advantage (Piwowar and Vision, 2013; Clayson et al., 2021), facilitates compliance with publishing guidelines (Nosek et al., 2015), increases job competitiveness (Nosek et al., 2022), expands networks (and, consequently, chances of collaboration), and more (McKiernan et al., 2016; Allen and Mehler, 2019). As such, these practices do require more upfront work, which may make them seem like they are slowing down the research process, but they pay off in the long run for the individual researcher as well as for the field. It is important however, to remember that the incentive structure needs to change so that the adoption of these practices becomes larger and its burden is not only the younger generation’s responsibility (Allen and Mehler, 2019).

2.3.1. Open Science Framework (OSF)—OSF is a general platform that supports collaboration and reproducibility by providing tools for efficient collaborative work and the coordination of different actors across a wide range of activities. For example, it allows researchers to manage, document, and share all the products of their workflow, possibly including the pre-registration of the initial idea, code repositories, and intermediary reports, up to the preprint of the final report. OSF is a free online platform developed by the non-profit organization Center for Open Science. This and other resources and general-purpose tools are listed in Appendix Table A1.

2.3.2. The Turing Way—The Turing Way (TTW, Community et al., 2019, also see Appendix Table A1) provides instructional resources on how to conduct reproducible and replicable work. TTW is an open-source, community-developed online collection of guides to reproducible, ethical, inclusive, and collaborative data science. Here we

highlight a selection of TTW guides. (1) *Guide for Reproducible Research* provides general recommendations for implementing reproducible research practices, such as storing and sharing the project's computational environment, using a version control system, and testing the code. (2) *Guide for Project Design* covers how to effectively plan and design a research project. For example, it illustrates how to write a roadmap of the project's goals, resources, and needed actions, how to structure a project's repository (e.g., folder organization), and how to clearly set up contribution pathways when working openly. (3) *Guide for Communication* offers recommendations for communicating research to wider audiences (e.g., through blog posts, podcasts, or social media), or to a specialized public (e.g., during posters and conferences talks), in an inclusive and accessible way. (4) *Guide for Collaboration* provides advice on how to set up remote collaborative projects, while working openly and encouraging diversity. In order to prevent human conflicts overshadowing collective endeavors, one of these recommendations is to include a Code of Conduct that explains how contributors are expected to behave (and what to do if these expectations are violated). (5) *Guide for Ethical Research* discusses the main concepts and institutions associated with ethical research, and provides examples through real-world case studies. These guides are continuously updated by the community, to accelerate the inclusion of new reproducibility practices and GSP.

2.3.3. #EEGManyLabs—The #EEGManyLabs project (Pavlov et al., 2021) is a large-scale, international replication effort. Concerns regarding the replicability of psychological phenomena have been spreading to multiple subfields of psychological science (Open Science Collaboration, 2015). The scale of the replication crisis in MEEG research has yet to be defined but, given the relatively noisy data, challenges to acquire data from large numbers of participants, and the analytical flexibility in MEEG analysis, MEEG research is unlikely to be free from replicability concerns. In response to this, the ongoing #EEGManyLabs project (Pavlov et al., 2021) aims to replicate over 20 of the most influential psychophysiological EEG findings, in at least three laboratories each with largely increased statistical power. To ensure high quality, each replication effort takes advantage of the Registered Reports format (see Section 2.2), uses standardized operating procedures and analysis pipelines, and passes internal review by the advisory board and original authors.

This extensive commitment to collaborative and open research is expected to increase collective confidence in EEG research, inspire new standards for reporting EEG findings, provide researchers with a large open database of raw data and analysis pipelines for further exploration, and an effect sizes catalog of various commonly studied EEG phenomena, including ERPs, in order to support the design and initiation of novel research. Besides the direct outcomes, the project aims to facilitate a shift towards collaborative psychological science and neuroscience. Studying increasingly subtle EEG effects requires large sample sizes, which are generally not easily achievable in a single lab. For EEG research to continue to thrive in the 21st century, multi-site high-powered studies are required, improving statistical power, reproducibility/replicability, and generalizability.

In addition to #EEGManyLabs, similar collaborative efforts have emerged in the past several years, for example, multi-site replication initiatives by Nieuwland et al. (2018), Nave et al. (2020), and Whiteford et al. (2020), or a recent Many Analysts initiative to investigate

effects of variability in the pre-processing and analysis pipeline (EEGManyPipelines). So far, they are still only sporadic, but may be the beginning of a more widespread trend. These initiatives and resources are listed in Appendix Table A1.

2.3.4. ERP CORE—The ERP CORE (Compendium of Open Resources and Experiments; Kappenman et al., 2021 see also Appendix Table A1) is a freely available online resource providing stimulus presentation scripts, data recordings, analysis pipelines, and results for 6 well-known event-related potential (ERP) paradigms. As such, it provides an example of a project that has successfully replicated influential ERP experiments and made all material freely available.

The replicability of MEEG research can be increased by taking advantage of information about optimal design and analysis provided by prior studies. However, ground truth is not typically known, making it difficult to know if the methods from a given prior study are actually optimal and generalizable. The ERP CORE was created to provide a reference point for future research by taking widely replicated MEEG paradigms, optimizing them, and providing information about them that can be used as solid ground for subsequent studies.

This open resource contains stimulus presentation scripts, data from 40 neurotypical individuals, and data analysis scripts for six common MEEG paradigms that are designed to isolate seven common ERP components: N170, mismatch negativity, N2pc, N400, P3b, lateralized readiness potential (LRP), and error-related negativity (ERN). These effects have been replicated so many times that the question is not whether they exist, but rather how best to obtain them. The specific versions of the paradigms in the ERP CORE were developed in collaboration with multiple experts with the goal of creating optimized versions that produce valid and reliable effects with approximately 10 minutes of data collection per task. Moreover, the online resource includes extensive information about data quality and effect sizes that can be used as a reference for comparison with new studies and to conduct power analyses. Finally, data-driven recommendations have been provided for time windows and electrode sites that are optimal for quantifying amplitudes of each component, which can be used to provide *a priori* analysis parameters for future research. This may be particularly useful for investigators acquiring MEEG data with low density/portable systems.

The individual ERP CORE tasks can simply be inserted without change into other studies (e.g., large-scale clinical studies that include an ERP measure) or used as a starting point for new tasks. The data processing pipelines can be used without change to provide *a priori* analysis methods or used as a starting point for new analyses. The relatively large existing dataset can be used to test new hypotheses (e.g., regarding correlations among components) or to assess how new analysis methods work across a variety of paradigms.

3. Collecting, analyzing, reporting, and sharing

Where the previous section considered planning and pre-planning stages, this section considers the actual execution of the work: the collection, analysis, and publication of data and results. For such work, there exists a category of GSP that is very specific to the exact procedures that are being performed, and the reader is encouraged to follow them where

appropriate. We will cite some of them here. However, as the main focus of this paper is on a more general form of GSP, this section will primarily contain broader considerations.

One key element of GSP is the understanding that hardly any decision with respect to a research project can be made without this decision affecting other aspects: experimental design, data collection and storage methods, signal processing steps, analysis pipelines, documentation, and reporting all interact with each other.

With respect to data collection, Section 3.1 discusses the availability of open acquisition protocols, considerations with respect to clinical recordings, the sharing of data, and the option of not, in fact, recording any data at all, depending on the needs of the project. Section 3.2 deals with software, providing help and resources both when writing code from scratch, and when choosing which MEEG software to use. Both newly written code and the choice of toolbox can have an impact on later steps in the project. Next, this section covers various aspects of data analysis: Signal processing (Section 3.3), statistics (Section 3.4), and machine learning (Section 3.5). Here, high-level GSP primarily emphasizes the general understanding of each method and its assumptions, especially considering its dependencies and interactions with respect to experimental design. Finally, Section 3.6 addresses standards surrounding data organization, analysis, and reporting. Making use, and contributing to the development, of such standards enable, among other things, easier sharing, collaboration, and (mass) analysis during a project, and accurate documentation and reporting afterwards.

3.1. Data collection

Data collection is often the next step after the project has been properly planned. Below, we describe GSP related to stimulus presentation, present open acquisition protocols for scientific studies, and introduce some relevant considerations for research studies in clinical settings. Of course, not every scientific project requires acquiring new data. Open repositories (where pre-recorded datasets can be downloaded free of charge) or carefully designed simulations can allow some projects to go without recording any data at all. Therefore, we will also present some examples of MEEG open data repositories and tools for data simulations. Thus, in general, GSP for data collection serves not just to record the cleanest possible data, but also to carefully consider what the needs of a project are, what data and procedures may already be openly available and verified, and which products of the current work may be of use to others.

3.1.1. Stimulus presentation and synchronization—In MEEG experiments, the main experimental manipulation is usually done in software, e.g., by presenting different stimuli to the participants. For stimulus presentation and response registration, a wide array of software solutions is available, including online platforms. In the context of GSP, it is important to be aware that even the same software may perform differently on different hardware with respect to a number of parameters, including latency and jitter. Therefore, the main recommendation here is to properly measure and document these parameters. This recommendation extends to all data that is simultaneously recorded, since any differences in timing between modalities may need to be corrected for during analysis. Bridges et al.

(2020) provide timing measurements for a number of stimulus presentation packages, but also stress the importance of making one's own measurements.

Besides stimuli, responses, and the MEEG itself, MEEG experiments may additionally record audio, video, motion capture, eye tracking, and any other modality deemed relevant. This can result in complex set-ups where each modality is recorded by a different device, at a different sampling rate, in a different format, and relying on a different internal clock. As a result, time stamps, latency, and jitter may vary for all of these modalities. Both for analysis and later reproducibility, it is important to measure these differences, and correct for them where possible. The three most common ways for measurement and/or synchronization are the following.

The most accurate method employs dedicated hardware capable of recording and/or comparing multiple sources of data simultaneously. For example, such a device can acquire physical measurements of visual and auditory stimuli using photodiodes and microphones, respectively, and compare their timings. Such physical measures may also be recorded directly by some MEEG amplifiers as auxiliary channels, allowing even jittered offsets to be corrected post hoc.

When no direct hardware connection can be made between data streams, one alternative is to introduce external triggers that can be recorded by all modalities separately. This creates one identifiable point in time along which modalities can be aligned. Note that, because clocks may drift, the procedure may need to be repeated at set intervals.

The lab streaming layer (LSL) (Stenner et al., 2021) (see Appendix Table A1) presents a software-only solution to synchronizing different modalities. This is a free and open source software framework that allows for the transmission and collection of any number of data streams across one or more devices on a local network. One of its features is automatic data synchronization, and an increasing number of manufacturers of electrophysiological hardware support LSL by providing real-time access to their data in this open format. However, LSL does not know the age of the sample when it is first recorded, or the delay between the hard- and software: therefore, at least an initial hardware-based measurement of this delay is still recommended.

Also note that calibration test results represent research data in their own right, and the results of teams using different combinations of hardware and software can provide valuable information about optimal configurations for certain kinds of tasks. This is particularly important in experimental protocols being implemented in multiple testing sites. To make optimal use of this information, software and hardware configurations need to also be documented in detail, as discussed in Section 3.6.

3.1.2. Open acquisition protocols and MEEG data collection—Open acquisition protocols provide an open-science approach to optimizing the procedures and methods around data collection. MEEG recordings contain brain activity of interest, but these signals are tiny and are mixed with a variety of biological and nonbiological noise sources that can dwarf the neural signals. Substantial effort and ingenuity are required to extract reliable

and meaningful brain signals from the noise. Open acquisition protocols help formulate and communicate some of the required procedures.

The most important step in dealing with noise is to record the cleanest possible data (Sinha et al., 2016). Post hoc solutions may exist, but the most straightforward signal processing method for reducing noise — filtering — can cause significant distortion of the signals in the time domain, shifting latencies and even producing artificial peaks and oscillations (Tanner et al., 2015; Widmann et al., 2015; Yael et al., 2018). Traditionally, some sources of noise in EEG recordings were reduced by abrading the top layer of the scalp and adding conductive gel or solution to improve the low-impedance contact between the electrode and the living tissue of the skin. Today many researchers use high-impedance systems for reasons of speed and safety. Unfortunately, these systems are prone to an increase in skin potentials and other sources of low-frequency noise, especially when the recording environment is warm, which can in turn reduce the statistical power of the study (Kappenman and Luck, 2010). Another important technical factor is an impedance mismatch between an active EEG recording sensor and the reference electrode, which increases the amplification of non-EEG signals. Many other factors also impact the noise level, such as muscle tension, movements, perspiration, and nearby electronic devices. Thus, we are trying to record a needle of brain activity in a haystack of noise.

To promote optimal methods for clean data recording, a group of researchers has published a detailed EEG recording protocol that provides a precise description of the “special sauce” that they developed over decades for maximizing the signal-to-noise ratio (SNR) (Farrens et al., 2020). For example, participants are asked to vigorously comb their scalps prior to electrode application, and the electrode wires are oriented during electrode application to minimize lead tension and movement artifacts. Details like these do not typically appear in Method sections of journal articles. A revised version of the protocol is also available with modifications designed for safe testing during the COVID-19 pandemic (Simmons and Luck, 2020). In addition to sharing these tried and tested methods, this protocol also serves as a model for other researchers who wish to publish their protocols. It is published on Protocol Exchange (see Appendix Table A1) — a free protocol repository that also provides an automatic way to create a digital object identifier (DOI) and indexing on Google Scholar. Other researchers are strongly encouraged to publish their recording protocols, which should increase transparency and reproducibility as well as spreading more effective recording methods.

In recent years, several scientific publications have provided guidelines for the collection of MEG data (Bagic et al., 2011; Gross et al., 2013; Hari et al., 2018). In the clinic, patients present their own confounds and contaminants that must be accounted for, discussed in detail in the next section. Recently, (Mosher and Funke, 2020) provided practical guidelines for the preparation of MEG instruments and patients or subjects for routine operations. Among other things, this acquisition protocol highlights localization as a crucial issue in multiple ways. Recording individual locations of the electrodes and, importantly, head position indicator coils, e.g. using a polhemus stylus, is crucial for accurate source localization. Also the “location” of the participant’s head as a whole needs special attention: it must be deep inside the helmet, and must not be allowed to “slump” out, lest proper

brain coverage is lost and unrecoverable. Other considerations for MEG data collection are covered in more detail in Mosher and Funke (2020).

3.1.3. The clinical setting and intracranial data collection—In the clinical setting, recordings collected from patients provide unique opportunities not only to diagnose disease but also to advance our understanding of healthy brain functioning. However, working with patients requires special considerations. For smooth clinical operations, researchers must respect the patient’s comfort, contemplate costs, and ensure the collection protocol is designed and followed properly; otherwise, not only is rescheduling the patient difficult, but the very act of rescheduling places undue stress on the patient, who may assume that some serious finding in their data has necessitated a second exam. For outpatients (i.e., patients not admitted to a hospital), the procedures otherwise follow the same GSP as for common acquisition protocols given above. For inpatients (i.e., hospitalized patients) additional considerations are required, particularly for patients with surgery for the implantation of invasive electrodes.

Intracranial EEG (iEEG) signals are collected for the investigation and treatment of various pathological conditions including epilepsy, Parkinson’s and tumors. Epilepsy, in particular, often requires the placement of dense arrays of electrodes subdurally (electrocorticography: ECoG) or intracerebrally (stereoelectroencephalography: SEEG) over diverse cortical regions. In contrast with non-invasive techniques, intracranial recordings present several advantages such as: (1) enhanced SNR, (2) focal recordings directly from the neuronal source, (3) minimal signal distortion due to the skull and other tissues (Fahimi Hnazaee et al., 2020), and (4) the possibility to record from deep brain structures (Buzsáki et al., 2012; Parvizi and Kastner, 2018). In addition to SEEG or ECoG, some centers add microwire electrodes to benefit from signals with higher spatial and temporal resolution that give a precious and rare access to multi and putative single neuron activity in humans. Microwires are very sensitive and their use requires special attention at all steps of the acquisition, from the implantation to the recording. For detailed recommendations on microelectrode manipulation and technical settings, see Lehongre et al. (this special issue).

Research with iEEG data provides a unique opportunity to address pathophysiological, physiological or cognitive questions, but because of their invasiveness, iEEG implantations are performed only with the strongest consideration for the standard of care for the patient. This requires carefully selected research protocols, i.e., protocols where iEEG has a significant and specific benefit relative to non-invasive techniques and are suitable to the patients’ abilities. The “inpatient” collection setup is technically very challenging, from obtaining patient consent (explaining the procedures) to data acquisition, and requires limited interference with clinical procedures. The recruitment of participants is therefore limited, requiring multiple months/years or the involvement of multiple centers in order to accumulate an adequate number of patients. For physiological/cognitive studies, the patient’s pathology and medication regimen can affect the recorded activity relative to a healthy participant. Cognitive impairments can make the task difficult, frustrating the patient and impairing the protocol. The brain coverage of iEEG electrodes is individualized, focal, and sparse, and thus can complicate group studies (for further details see Dubarry et al., in the same issue).

Therefore, a close collaboration between clinical staff and researchers is key to collecting high-quality iEEG data. Optimally, clinicians themselves are involved in the research protocols, and staff are dedicated to interfacing between clinicians and researchers (Mercier et al., in the same issue).

Simultaneous recordings of iEEG, scalp EEG, and/or MEG capitalize on the strengths of each respective technique to describe the underlying brain activity accurately (high SNR) and comprehensively (overall view). First, from a clinical perspective, simultaneous recordings of iEEG with EEG and/or MEG can provide complementary information regarding the epileptogenic zone (Santiuste et al., 2008; Kakisaka et al., 2012; Gavaret et al., 2016). Second, in basic research, iEEG has sometimes been considered as capturing the “ground truth”, thus providing methodological advances for evaluating how known cortical activity is captured by non-invasive methods (Koessler et al., 2015; Pizzo et al., 2019; Seeber et al., 2019). Finally, in systems and cognitive neuroscience research, the analysis of signal fluctuations evoked by individual stimuli in depth and surface recordings can be exploited as a crucial source of information for refining our understanding of the neural activity underlying specific cognitive processes (Dalal et al., 2009; Dubarry et al., 2014). Thus, understanding the relationship between surface signals and the spatiotemporal configuration of the underlying cortical sources is central to both basic fundamental and clinical research.

However, introducing scalp EEG and/or MEG simultaneous recordings to an already highly-constrained recording setup for iEEG creates a large number of important considerations. For scalp EEG, the presence of the surgical dressings can be challenging, and the set-up may require constant regelling of scalp electrodes and addition of physiological adhesive to ensure that scalp electrodes remain in place and operational. Signal analysis must also take into account the dramatic changes in electric field caused by the implants, their burrholes and craniotomies, which create large skull discontinuities (Dalal et al., 2009; Kirchberger et al., 1998). For MEG, conducting simultaneous recordings with iEEG requires a dedicated MEG facility housed in a clinical setting. The complex setup presents obvious physical constraints (e.g., fitting bulky iEEG electrode connectors within the MEG dewar) and significant technical challenges that occur when introducing metallic materials into a shielded MEG environment. Because the patient is away from the clinical monitoring unit, the exam must be carried out within a limited period of time, usually not more than one hour. Furthermore, MEG signal quality can be drastically affected by the presence of the iEEG electrodes and equipment (e.g., cables, connectors). To date there is only one group that has achieved the simultaneous acquisition of MEG, EEG and iEEG signals (Dubarry et al., 2014; Gavaret et al., 2016).

In summary, simultaneous intracranial and scalp recording sessions require complex logistics and organization between clinical staff, patient, technologists/research engineers, and clinicians to ensure smoothness of procedures and minimization of time spent outside of the clinical monitoring unit (see also Mercier et al., in the same issue). Planning such recordings should therefore only be considered when clearly motivated by the research hypothesis and strongly supported by the clinical staff. Nonetheless, recordings in the

clinical setting provide a unique opportunity to research various aspects of brain activity, with potentially significant benefits to understanding the human brain.

3.1.4. Open MEEG data repositories—Data sharing is important for solid scientific progress (Gorgolewski and Poldrack, 2016; Poldrack et al., 2017, 2020). While significant resources have been invested in neuroimaging studies worldwide, the number of projects which publicly share their data remains limited. Over the last years, some initiatives have emerged to openly share MEEG datasets. These repositories foster reproducibility of results and help to answer specific explorative scientific questions not originally conceived at the time of data collection and study design, increasing data usage and longevity. Furthermore, they allow data aggregation, attaining larger sample sizes which could be used to increase statistical power and for approaches involving artificial intelligence (see Section 3.5). Open repositories also facilitate data access to researchers with less funding opportunities and resources, giving back to society part of the value invested. Some examples of open MEEG repositories are: Human Connectome Project (HCP) (Larson-Prior et al., 2013), The Open MEG Archive (OMEGA) (Niso et al., 2016), The Cambridge center for Ageing and Neuroscience (Cam-CAN) (Shafto et al., 2014; Taylor et al., 2015), The Temple University Hospital EEG Data Corpus (TUH EEG Corpus) (Obeid and Picone, 2016), and OpenNeuro (Markiewicz et al., 2021). See more resources in Appendix Table A1.

3.1.5. Simulating data—Up to this point we have discussed data obtained by recording brain activity from actual (human) participants or patients. However, there are a number of reasons why the generation of so-called synthetic, toy, or simulated data may sometimes be preferable to data from live participants.

One practical reason concerns the cost of data collection. Recording real data requires the development and validation of specific experimental paradigms, time to record from a sufficient number of participants, and the acquisition of all necessary supportive material including human resources. Simulating data can be much more efficient, especially given the dedicated simulation tools available today. A combination may also be useful, for example when testing planned analysis pipelines on simulated data prior to implementing actual data collection.

Another reason why it may be preferable or even necessary to use simulated data instead of real data is that only simulated data can provide a fully known ground truth. To verify the accuracy of a method, its results need to be compared to a known truth. Unfortunately, with MEEG recordings no such ground-truth ideal of brain activity is available. Therefore, researchers often construct synthetic data in such a way that its ground truth is known, allowing the idiosyncrasies of a method to be properly evaluated.

A number of software packages have recently become available for creating simulated data using a number of fundamentally different methods. For example, generative-adversarial neural networks can mimic existing EEG data (Hartmann et al., 2018). Such neural networks take previously recorded real data and generate new synthetic data that cannot be distinguished from the original data it is based on. As such, the resulting data does not contain a known ground truth, but can be used for data augmentation, data recovery, or

up-sampling. The Virtual Brain, on the other hand, uses interconnected neural mass models based on differential equations to simulate the low-level dynamics of clusters of neurons (Sanz Leon et al., 2013). This results in a potentially highly detailed ground truth. However, the most common method of MEEG data simulation employed in the past decades relies on a simplified forward model which describes MEEG as a combination at the scalp of activity projected from a limited number of generally independent sources in the brain. Following this approach, the simulation of MEEG data requires the generation of a number of source activations, the linear mixing of these source activations to obtain scalp activations, and the optional addition of noise. By generating these three factors themselves, scientists can be sure to know the ground truth of the data. Previously, researchers often wrote their own code for such simulations, limiting standardization and reproducibility. However, one recent free and open-source MATLAB-based EEG data simulation toolbox, SEREEGA, is based on this common method and provides a standardized, reproducible framework to simulate data in this fashion (Krol et al., 2018). See Appendix Table A1 for a selection of these and other MEEG simulation tools.

Note that any simulation approach is, itself, a method that makes certain assumptions. As such, it is a tool that provides certain functionality, but cannot guarantee any particular quality or appropriateness of the outcome. Each simulation needs to be performed for a particular purpose, requires different parameters, and makes specific demands with respect to different kinds of validity. As is the case with all tools and methods, it remains crucial to carefully consider a simulation's parameters and their appropriateness for the particular task at hand. For example, when a method that assumes linearity in the data performs accurately on simulated data that has been constructed using a linear model, it has not been proven that this method will transfer well to real data, which may not exhibit the same linear characteristics. That said, simulation is a powerful tool, as it gives full control over the data. An appropriate simulation can provide a standardized way to objectively and reproducibly validate and evaluate the tools we develop and use.

3.2. Software strategies for GSP

Computer software is at the core of any MEEG data analysis project. Be it code that scientists write from scratch, or software toolboxes that support them, the main purpose of GSP at this level is generally to promote its reliability. However, correct and readable code plays a key role in reproducibility as well. A further set of GSP is relevant when using other, published software. In this section, we first highlight a few key principles that scientists can benefit from when designing code, and then briefly review how today's most widely used software toolboxes have developed and engage in GSP. In general, GSP for the proper writing and use of software can help provide clarity and transparency to the research process, aid reproducibility, and verify the robustness of results.

3.2.1. Designing code for correctness and reproducibility—A core ingredient of a scientific experiment is to document, in detail, what steps have been performed, so that they may be reproduced. When the number of steps is small, the Methods section of an article may contain them in full. However, journal page limits ensure that this is frequently not the case in modern MEEG studies, requiring Methods to be summaries that omit many

important details so as to not cloud the ‘big picture’. This leaves the programming code itself as the authoritative source of the exact analysis procedures used in the collected data. With modern programming languages and tools, code can serve this purpose well, if *it is consciously written to do so*. If, however, the programmer’s focus is solely on obtaining the desired output, without regard for transparency in how results are obtained (e.g., documentation, easy to follow program flow, version control), chances are high that the results will not be reproducible, or maybe even wrong (Miller, 2006; Casadevall et al., 2014; Pavlov et al., 2021).

Much has been written already on the topic of how to produce code that clearly communicates analysis steps, and is applicable to stimulus presentation and experimental control scripts. The reader is encouraged to look at the writings of van Vliet (2019) and Wilson et al. (2014). The most important thing is to have the mindset that scientific code *is not done until it is well organized*. This has little to do with technical skills. Rather, this is about proper “housekeeping”: making sure unneeded code and files are archived and removed from the main project, making sure everything is named properly, long and convoluted scripts are cut up into commented modules that are easier for others to understand, etc. As the analysis advances and new code is written, clutter inevitably accumulates. Fortunately, there are tools for organizing files containing code, such as version control systems (see Appendix Table A1). With some discipline, these tools allow keeping track of every change made to the code, and allows efficient collaboration at scale. Furthermore, more and more analysis toolboxes such as those presented below offer script generators to improve reproducibility and generate code that is easy to understand and adapt (Es et al., 2021), which reduces the burden on the researcher.

3.2.2. Software toolboxes for MEEG data analysis—A number of comprehensive software toolboxes exist for analyzing MEEG data. When considering reproducibility, the choice of an analysis environment to work in should not be made lightly, as this decision may commit the data analyst for not only the current project but also future studies. Regardless, some points to consider are the required (and available) level of programming skill and financial resources, the availability of a graphical user interface, and the availability of local software support and expertise. However, ultimately, not all toolboxes are equally well-suited for all types of analyses or modalities (EEG, MEG, fMRI...). Although the most common analysis methods are implemented in most toolboxes, some offer more advanced algorithms in one or another environment, and more advanced projects may require applying methods available across several toolboxes. In that case, care should be taken in assuring that all technical information is handled properly across environments. Monitoring and seeking advice from the community via the relevant user forums is advisable.

Table 1 highlights current differences and summarizes the most relevant features of five of the most commonly used MEEG analysis environments to date. These toolboxes are under constant and active development, and the developers cooperate to reimplement each other’s methods and reflect on the best ways to do so (e.g., Jaiswal et al., 2020, or Delorme et al., 2022). These toolboxes all implement standard and advanced analyses, and enforce GSP in some way, but each with their own peculiarities, inherited from the laboratory traditions, equipment, software environments and applications for which they were developed (see Box

3). Attention to these specifics may aid the decision for which toolbox(es) to optimally use for a given project.

In line with current GSP (see Section 3.2.1), these toolboxes can generate scripts and reports. In particular, reports that mix code and human readable narratives are handy tools that allow saving the results and the code that generated them in one document, easily accessible, and readable with any document viewer. If these reports expose all (including default) parameters and software versions, they can readily be used to reimplement an analysis with minimal burden, even in a relatively distant future. This feature is particularly important because backward compatibility with previous versions can hardly be ensured. At the time of publication, archiving of code and reports on to a public repository, ideally with a permanent digital object identifier (DOI), is a desirable step that greatly facilitates reproducibility.

Given that each toolbox is likely to have its own idiosyncrasies and unique implementation even for ‘standard’ methods, it is GSP to perform the same analysis pipeline in different toolboxes and compare their outcomes in order to avoid results being affected by such issues. When the same analyses across toolboxes support the same conclusion, this strengthens the robustness of the results.

3.3. Signal processing

Signal processing is the step in the research lifecycle that extracts from the raw data the key information needed to answer the scientific question being asked. Previously published GSP primarily concerns the nature of the pitfalls or limitations with respect to various methods, and the recommendations to avoid them. However, these vary largely depending on the specifics of the methods and characteristics of the recordings. Therefore, given the ever-growing number of tools routinely used to (pre)process MEEG data and the wide range of experimental protocols in the field, a systematic exploration and exhaustive account of GSP for all MEEG signal processing is not practically feasible (although a number of guideline documents exist, listed in Section 1). Instead, this section offers general high-level recommendations for better planning and designing the signal (pre)processing components of an MEEG study.

First, the goal of the study (i.e., the question it seeks to resolve) dictates a possible set of signal processing methods. From this set, the analysis methods should be chosen *before* finalizing the experimental task design (e.g., see Section 2.2). This ensures that collected data can be optimally examined with the projected tools. For example, if the plan is to examine slow oscillatory components in the signal, or run detrended fluctuation analysis on the data, it is crucial to make sure the temporal windows that will be analyzed are long enough to allow for such analyses, or that the data themselves are collected using appropriate analog filtering. Although it is possible to consider one’s analysis options based on the properties of the available data, this approach might result in “p-hacking” (Simmons et al., 2011) and other biases mentioned in previous sections.

Second, it is very important to understand how different signal processing tools transform the data, and what the transformed signal reflects or measures. This does not mean

that everyone should master the mathematical and implementation details of the signal processing tools they use, but it is important to develop an *intuitive understanding* of these tools and to be aware of the parameters and default settings associated with the various processing steps in the pipeline (e.g., what filter parameters were used and why). The minimal level of understanding of a tool should include an appreciation of its limitations or pitfalls. A case in point is the importance of understanding the limitations and specificities of distinct connectivity metrics. For example, problems associated with metrics sensitive to zero-phase lag coupling, and the potential limitations that arise with methods thought to overcome these very issues, can be understood without formal training in mathematics, but this is not obvious from the outset and requires practice (Palva et al., 2018).

Moreover, an issue that is sometimes overlooked is the feasibility of connecting various processing steps into one signal processing pipeline, and the order in which multiple operations are performed. Simple examples include the importance of computing time-frequency maps in single-trial data before averaging across trials when investigating induced responses (Tallon-Baudry and Bertrand, 1999). Likewise, it is important to verify whether certain pre-processing steps (down sampling, filtering, ICA etc.) preclude the subsequent application of specific signal processing tools (Cohen, 2014).

Third, when choosing certain methods, it is important to consider implicit assumptions about data properties and whether these are warranted. As an example, choosing to only compute power in a given frequency band (e.g., theta-band) without examining the full power spectrum is usually based on the assumption that the data actually exhibit specific oscillations at that frequency. This assumption can be confirmed by examining the entire spectrum for a peak at the frequency range of interest. As a matter of fact, several methods can verify the presence of oscillations and distinguish them from the background 1/f component of MEEG data, for example using Empirical Mode Decomposition (EMD, Quinn et al., 2021), Fitting Oscillations and One-Over-F (FOOOF, Donoghue et al., 2020) or Irregular Resampling Auto-Spectral Analysis (IRASA, Wen and Liu, 2016), as reviewed by Gerster et al. (2021).

Fourth, similar to replicating results using different toolboxes (as mentioned in Section 3.2.2), conducting the same analysis using alternative signal processing methods can be a useful approach to test for robustness of an observed effect and potentially to troubleshoot analyses. Obviously, attempting to replicate (or compare) findings using multiple methods requires time and effort, and it cannot be done for all analysis steps. However, it can be a useful practice that can be applied to a few key steps in the pipeline. Examples include using more than one source estimation technique (or more than one implementations of the same technique, cf. Westner et al., 2022), or using both wavelet-based and Hilbert-based implementations of spectral analyses metrics. Likewise, exploring a range of values for the parameters of a given analysis rather than sticking to default values can be a way to probe the robustness of the results as a function of parameter settings. If small changes in the parameters lead to substantial changes in the results, one should not systematically dismiss the results as unreliable, but rather view this as a warning call for further investigation.

A fifth, generally important habit to develop in connection with signal processing is frequent data visualization. As often as possible, the output of intermediate analysis steps throughout the signal processing pipeline should be plotted and carefully examined. This can be achieved by using modular coding approaches with opportunities to examine the output of each signal processing stage (see also Section 3.2.1). This also facilitates rapid trouble-shooting and in some cases might even trigger new ideas for data analyses.

Finally, the recent proliferation of studies using machine learning (ML; e.g., brain decoding) has generated a lot of interest and excitement in MEEG. However, the rise of “data-driven neuroscience” should not occlude the fact that assumptions made about the data and the implicit knowledge of the experimental paradigm both dictate analysis choices and/or parameter tuning (e.g., choice of extracted features, task design, or target classes in decoding studies using supervised learning, etc.). Data-driven approaches are not more “objective” methods, and their rise is not a sign that hypothesis-driven neuroscience has come to an end: both approaches complement each other. The best practices and pitfalls of MEEG signal processing in the context of ML analytics are largely the same as those that apply when exploring data from other domains. Among other things, one should watch out for violations of the strict separation between training and test data (e.g., avoid feature preprocessing that induces dependencies across the whole dataset) Section 3.5. presents more recommendations for the use of machine learning for MEEG data analysis.

3.4. Statistics

In experimental MEEG data analysis, statistics are mostly considered as an inference tool. Given a particular experimental design, statistics help experimenters decide whether or not a given treatment affects measurements, or if two groups have significantly different signals in a given region of interest (electrode, time window, frequency band, anatomical location...). It is one of the challenges of GSP to keep these inferential statistics valid across the various situations in which they are used. Relevant topics are extensively covered in textbooks on statistics. Here, we address GSP with respect to four more general statistical aspects of emerging MEEG analyses.

3.4.1. Data-driven and permutation-based analysis—Over the last few decades, statistics have become a tool for data-mining in exploratory (or so-called data-driven) analysis and whole-brain inference, where mass univariate statistics (Groppe et al., 2011) and multivariate analysis (Cichy et al., 2014; King and Dehaene, 2014) deal with the challenges of multidimensional MEEG data. These new use cases have stimulated a number of changes in GSP. A well-known example is the adaptation of decision thresholds to the number of tests being performed (the so-called multiple comparisons problem). Performing tests across the entire data space increases the chances of reporting false positives. Methods such as family-wise error rate (FWER) or false discovery rate (FDR) prevent this type of inflation. Among those methods, non-parametric permutation-based approaches (Maris and Oostenveld, 2007; Smith and Nichols, 2009; Sassenhagen and Draschkow, 2019) currently outperform other methods (Hayasaka and Nichols, 2003; Nichols and Hayasaka, 2003; Puoliväli et al., 2020). For decoding approaches and out-of-sample generalizability, permutations have also been used for a data-driven robust estimation of “chance level”

(Combrisson and Jerbi, 2015), adapting for choices of cross-validation schemes (i.e., how to split testing and training sets) and decoders. For a comprehensive introduction to non-parametric (permutation-based) statistical testing in MEEG, the reader is referred to Maris and Oostenveld (2007).

3.4.2. Simulations for prospective power analysis—It was recently argued that neuroscience studies are often underpowered (Ioannidis, 2005; Button et al., 2013; Szucs and Ioannidis, 2017), resulting in inflated and unreplicable effect sizes. It may also waste resources, as experiments are being conducted without ensuring that proper power can be achieved. To avoid this, actual effect sizes are currently being measured and registered with projects such as ERP CORE, or #EEGManyLabs (see Section 2.3.3). In order to evaluate the detectability of an effect of a particular size in a given experiment, simulations are important. As already noted (Section 3.1), simulations can generate data containing effects and noise of known amplitudes, making it possible to assess statistical power as a function of e.g. sample size. In a recent simulation-based study, Chaumon et al. (2021, this issue) examined how the expected spatial properties of sources of MEG activity affect statistical power. This study clearly showed that spatial variability in the source of the signal and the type of contrast measurements strongly affect statistical power. Thus, taking into account individual anatomical variability of expected active regions in a given dataset is key in experimental MEEG studies, since the number of samples (trials and participants) that are required to achieve a given level of statistical power vary several fold between areas with lower spatial variability (e.g., in the precentral sulcus) compared to regions with higher spatial variability (e.g., lateral occipital cortex). Specific MEEG simulation toolboxes (Krol et al., 2018), or dedicated statistics simulation toolboxes (Lakens and Caldwell, 2021) can be used for prospective power analysis. If hypothesis testing is conducted under a Bayesian approach, simulations can help set a maximum sample size for a Sequential Bayes Factor design. In this design, the researchers run their analysis pipeline after data from each participant (or batch of several participants) are collected. Data collection ceases when there is strong evidence for either H1 or H0, or when the maximum sample size has been reached (Schönbrodt and Wagenmakers, 2018).

3.4.3. Quantifying data quality—A fundamental challenge in MEEG research – where signals are typically tiny relative to noise – is to obtain the highest-quality data, i.e., precise, reliable measurements of the brain signal of interest. Remarkably, however, the field does not have a standard and widely-used approach to quantify data quality in a given study. In most published papers, there is simply no way to objectively evaluate how “clean” the data are. Two approaches have recently been developed to address this issue in the context of averaged event-related electrical activity (but could easily be applied to magnetic data).

One of these approaches takes traditional metrics of “reliability” from psychometrics (Rust and Golombok, 2014) and applies these metrics to ERP amplitude or latency measures (e.g., Olvet and Hajcak, 2009; Pontifex et al., 2010). In psychometrics, reliability is typically defined as the proportion of the total variance across participants that is the “true score variance” (true differences between participants, as opposed to differences due to noise in the data). This approach is particularly useful in individual differences research, which asks

how well a neural measure correlates with some other measure (e.g., with a symptom score). The correlation between two variables is straightforwardly limited by the psychometric reliability of the individual variables. Thus, an MEEG measure must be reasonably reliable to be used in correlational analyses. This approach has two limitations, however (see Luck et al. 2021). First, it provides a single reliability value for an entire group of participants and provides no information about the data quality for individual participants. Second, the reliability value depends on the amount of true score variance present in the group of participants, making it difficult to generalize across subject populations.

The second approach to quantifying ERP data quality is called the *standardized measurement error* (SME) (Luck et al., 2021), and it is an extension of the general concept of the standard error of measurement. The SME value for a given participant quantifies the precision of an ERP amplitude or latency score for that participant (i.e., the extent to which you would expect to obtain a similar score if you repeated the experiment multiple times for that participant). Unlike psychometric reliability values, the SME is computed individually for each participant. However, SME values can be aggregated across participants to quantify the overall data quality of a given experiment and to estimate the impact of noise on the effect size and statistical power of the experiment. The SME can also be combined with a new power calculator that takes into account the number of trials as well as the number of participants (Baker et al., 2020), making it possible to predict the effects of changing the number of trials in future experiments with the same basic paradigm. A basic version of the SME is now automatically provided when averaged ERPs are created by ERPLAB Toolbox (an EEGLAB plugin; Lopez-Calderon and Luck, 2014), and example code is provided for more complex applications (Stewart and Luck, 2020). An important limitation of the SME is that it is limited to amplitude and latency measures obtained from averaged waveforms and cannot be applied to single-trial analyses.

3.4.4. Multiverse analysis—Section 2.2 already mentioned the “garden of forking paths”, i.e. the almost innumerable possible variations in recording, pre-processing, and analysis pipelines (Gelman and Loken, 2013). This has, for example, recently been documented in a case analysis of the N400 ERP component (Šoški et al., 2021) (see also Section 3.6.2.). Of 132 analyzed N400 papers, each reported an idiosyncratic approach to acquisition, pre-processing, and analysis, with some methodological decisions being so diverse that almost no two studies took the same approach. The results from any one specific pipeline could be misleading: for example, it has been shown that sample size (Boudewyn et al., 2018), EEG recording systems (Melnik et al., 2017), electrode impedance (Kappenman and Luck, 2010), filters (Tanner et al., 2015), statistical analysis (Luck and Gaspelin, 2017), and other methodological decisions (Šoški et al., 2019; Sandre et al., 2020) can all affect study outcomes.

One of the approaches to addressing this issue (Hoffmann et al., 2021) which has recently been put forward is *multiverse analysis*, which involves applying a variety of pre-processing and analysis pipelines to a given dataset and comparing the results of all pipelines, allowing a more complete and a better theorizing of the constructs of interest (Steenen et al., 2016). Recently, this approach was implemented in a few methodological studies of ERP

components (Šoški et al., 2019; Sandre et al., 2020; Clayson et al., 2021), but this GSP can help as well in studies which are not primarily concerned with methodology issues.

3.5. Machine learning for MEEG data analysis

Artificial Intelligence (AI) is an ever-growing, transdisciplinary field impinging on many disciplines, including neuroscience. AI has produced some impactful applications in the MEEG field (Cichy et al., 2014; Hassabis et al., 2017; Lotte et al., 2018). This is particularly true for one of its emerging subfields (Bolt et al., 2021), machine learning (ML): see, for example, Lemm et al. (2011) for an introduction to ML for brain imaging. However, its transdisciplinarity combined with the complexity of physiological signals, inevitably could lead to important misinterpretations (Wang, 2019). Therefore, we here focus on a number of general issues and GSP with respect to ML as a tool for analyzing MEEG data.

ML generally employs *data-driven* algorithms which find patterns in data without any expert input – or indeed, without any constraints. These algorithms thus optimize their outcome regardless of the data's (neurophysiological) nature, leading to possibly unpredictable behavior (Hedderich and Eickhoff, 2021; Roberts et al., 2021). For example, unreliable or even false outcomes may arise in experimentally noise-contaminated data, as the algorithm may find noise to be a stronger predictor than cortical activity. This is because noise in MEEG data is not always random: it may be systematically correlated to the conditions of interest, leading the algorithm to use it to improve its performance outcome. Such artifact-based outcomes may seem reliable, but turn out to be meaningless for the neurophysiological effect one is trying to characterize.

Conversely, ML-based methods can produce unstable outcomes when they do not have enough data. Notably, deep learning-based MEEG analyses (Roy et al., 2019; Zhang et al., 2021) require a substantial amount of data to ensure powerful analyses, which is hardly met by the current datasets available (Hinss et al., 2021). When increasing sample sizes, care must also be taken to increase diversity in terms of participants, environments, and recording systems, in order to prevent the so-called “algorithmic injustice” (Birhane, 2021) which may otherwise result from data-driven methods.

While the availability and ease of use of ML algorithms is on the rise, e.g., through the Scikit-learn toolbox and resources (Pedregosa et al., 2011), the products of these algorithms often remain opaque. Their black-box nature makes it highly likely that their results are not interpreted correctly due to a poor understanding of the origin of their outcomes. Certain mathematical methods or functions, e.g., backward models in multivariate classification, have low to no neurophysiological interpretability and can lead to misinterpretations regarding the spatiotemporal origin of the neural signals of interest (Haufe et al., 2014). To ensure that the results are reliable, researchers should verify that algorithmic outcomes are neurophysiologically interpretable.

The above issues stress why signal (pre-)processing pipelines should also be used with a good understanding of their basic assumptions and effects on data so as to keep ML-based algorithm prerequisites fulfilled. As an example, and a case-study to be avoided in the Brain-Computer Interface community (Roy et al., 2019), there are algorithms reaching very

high classification rates thanks not to brain activity, but to muscle and/or ocular artifacts present in the same data. Here, experimental bias as well as (inappropriate) data processing and poor evaluation of algorithm outcomes lead to uninterpretable results in terms of brain activity. Aside from signal processing, experimental design may also need to be adapted for ML algorithms. For example, dataset slicing (mentioned in Section 3.3.) requires some conditions to be met (e.g., independence kept between the training and testing sets). Yet, some paradigms (e.g., block designs Lemm et al. (2011)) or data processing techniques introduce dependency between samples or across the whole dataset, thus violating these assumptions.

One project that deals with the above-mentioned issues is the Mother Of All BCI Benchmarks (MOABB; see Appendix Table A1), which intends to promote valid and reproducible BCI research through algorithm availability, benchmarking, ranking, and freely available datasets.

3.6. Current efforts for data standardization, analysis, and reporting

As mentioned in the introduction, one aim of this paper is to foster a general understanding of GSP that goes beyond previously established best practices. However, even general principles are sometimes best aided using specific tools. Here we discuss a small selection of current standards that provide a more specific, hands-on approach specifically for data organization and methods reporting, which have become more important in this era of open science.

Data storage and organization is currently largely performed in an idiosyncratic manner, sometimes differently even within the same lab. It is GSP, however, to organize all data in a consistent and structured way following the FAIR guiding principles (Wilkinson et al., 2016), where data should be Findable, Accessible, Interoperable and Reusable. This allows different analysis methods to have standardized access to it, while sharing and documenting it becomes easier and less prone to errors. Importantly, it also benefits the experimenters themselves when they revisit their data at later points in time.

Even with access to the same data, it may not be possible to fully reconstruct a study, as journal articles generally do not include complete and detailed MEEG methods descriptions (Clayson et al., 2019; Šoški et al., 2021). This would not be a pressing issue if results were robust to variations in the recording, pre-processing, and analysis pipeline. However, as mentioned in Section 3.4.4, this is not the case.

It is for these reasons that several initiatives have emerged in the neuroimaging community to deal with dataset organization, description, and reporting, three of which are introduced in this section. In the spirit of open science, these new initiatives have taken the form of living documents to keep up with ever-evolving perspectives of GSP, as well as the development of new tools and analyses. Additionally, these tools all make use of the more recent option to provide rich supplementary materials to help researchers achieve the level of clarity that cannot easily be met through methods descriptions in journal articles.

3.6.1. Data organization: BIDS—The Brain Imaging Data Structure (BIDS) is a community-led standard for organizing, describing, and sharing neuroimaging data (Gorgolewski et al., 2016). This standard facilitates data sharing and the development of analysis tools in the neuroimaging community. As an evolving standard, BIDS already supports multiple neuroimaging modalities including MRI (Gorgolewski et al., 2016), MEG (Niso et al., 2018), EEG (Pernet et al., 2019), iEEG (Holdgraf et al., 2019), and PET (Norgaard et al., 2021). The BIDS specification includes concrete details for folder and file naming, the choice of data formats, and the representation of metadata to be both human and machine-readable. Multiple accompanying data examples, and tools make it easy for researchers to incorporate BIDS into their current workflows, maximizing reproducibility, enabling effective data sharing, and ultimately supporting good data management practices (see Appendix Table A1). In particular, there are numerous tools to deal with MEEG data: (1) *BIDS converters* included in MEEG analysis packages; (2) *general tools for data querying and related operations*, e.g., PyBIDS (Yarkoni et al., 2019) and BIDS-MATLAB (Gau et al., 2022); and (3) *BIDS analysis tools*, e.g., BIDS Apps, which are containerized analysis pipelines that take BIDS-formatted datasets as their input and produce derivative data (Gorgolewski et al., 2017). Presently, the BIDS Derivatives specifications are under active development, setting the principles for organizing and describing outputs of processing pipelines, thus broadening the BIDS standard beyond “raw” neuroimaging data. BIDS is an open and inclusive community, and new members may start contributing to the initiative through the BIDS Starter Kit (see Appendix Table A1 for links to useful resources).

3.6.2. Templates for reporting: ARTEM-IS—As described above, achieving the level of detail sufficient to adequately replicate an MEEG study has proven to be a challenging task. Metadata templates that require precise numerical/categorical data to be filled can help to reduce error. Not only is template metadata clearer than verbal descriptions, but the standardized data format enhances searchability and simplicity for future metascience, data sharing, and reuse (Gau et al., 2019; Styles et al., 2021).

One such template is ARTEM-IS: an Agreed Reporting Template for EEG Methodology - International Standard (Šoški et al., 2020; Styles et al., 2021). Its content and structure is based on an analysis of social and technical challenges in methods reporting, with special attention to errors and omissions in the existing literature, as identified through systematic reviews. For example, in the Šoški et al. (2021) analysis of ERP methodology reporting, 100% of papers contained at least some omissions and 46% of papers contained at least some ambiguities in their described procedures. The data items identified in this review were used to create an evidence-based ARTEM-IS template customized for ERP research, that breaks down error-prone reporting items into smaller, clearer questions in a branching structure. By its design, this template requires data to be entered in a particular format, thereby improving the accuracy of documentation and standardizing the format of the reported details. The result is a living, version-controlled document. The pilot version of this template presented at the 2020 LiveMEEG conference consisted of 93 fields designed to capture ERP methodology up to the point of statistical hypothesis testing, in a loosely formatted spreadsheet (Styles et al., 2020). Through community effort, at the time of writing

in 2022 the template for methodology has been greatly refined and is currently being integrated into a structured format making up the backend of a webapp based on the model of COBIDAS Guidelines Checklist (see the section below and Appendix Table A1).

Inspired by the learned lessons of checklist adoption in aviation and surgery (Styles et al., 2021), ARTEM-IS also includes a Statement, which is a call to action, describing how if the community of MEEG practitioners want to: (1) improve clarity; (2) improve accuracy; (3) enhance documentation; and (4) deliver broad benefits, (5) community effort will be necessary. A link to the complete ARTEM-IS statement can be found in Appendix Table A1. Signatories to the statement are invited to participate in the further development of the ARTEM-IS template, and other similar initiatives in all fields of science are welcome to adopt the principles outlined in the ARTEM-IS Statement.

3.6.3. Data analysis and sharing: COBIDAS MEEG—As already noted, MEEG communities have proactively championed GSP, with specialist society journals being the most common format for sharing them (e.g., Donchin et al., 1977; Pivik et al., 1993; Picton et al., 2000; Duncan et al., 2009; Gross et al., 2013; Kane et al., 2017; Hari et al., 2018). The Organization for Human Brain Mapping (OHBM) has developed GSP white papers in neuroimaging, which are collectively referred to as COBIDAS – an abbreviation for the “Committee on Best Practice in Data Analysis and Sharing”. There is a version for MRI-based methods (Nichols et al., 2017) and for MEEG (Pernet et al., 2018, 2020).

COBIDAS MEEG guidelines first discuss pitfalls for data acquisition, analysis, and sharing for resting state and task-related studies. In general, the COBIDAS Committee believes that guidelines should not be prescriptive, as data (pre)processing pipelines vary by analysis method. However, the guidelines do include a set of Tables in the Appendix that list parameters that should be reported when preparing manuscripts and grants. Experimenters can pick and choose which of these tables are relevant at a given time. An used-friendly, Python-based checklist for MEG or EEG studies supports this process. This is a complement to an already existing eCOBIDAS App for MRI-based data (see Appendix Table A1). Importantly, COBIDAS MEEG is consistent with other neurophysiological guidelines, e.g., International Federation for Clinical Neurophysiology (IFCN) (Kane et al., 2017), and uses BIDS-consistent terminology.

COBIDAS MEEG guidelines differ in 3 main ways from existing MEEG guidelines. First, they specifically include practices of reproducibility and data sharing. Second, they were prepared as a living document with 2 branches: (1) a WordPress blog for feedback and comments (see Appendix Table A1), which remains open for the next incarnation of the guidelines; (2) an Open Science Framework version-controlled White paper (Pernet et al., 2018). Third, the target population is much broader than for previous guidelines and now also includes not only neuroimagers/neurophysiologists, but also the hardware and software engineers and physicists who write MEEG papers, theses, grants, and prepare Registered Reports and clinical trials, as well as those who review and evaluate them.

Some problematic issues discussed in COBIDAS MEEG are: (i) omitted critical data acquisition details (see also Šoški et al. 2021) from manuscripts and grants; (ii) the

distortion issue if data are acquired with a physically-linked ear or mastoid reference; (iii) inadequate treatment of statistical power and related issues, e.g., minimal effect size estimation for features of interest, use of independent data from existing literature and/or pilot data for choosing regions/sensors of interest; (iv) protocols for rejecting artifactual trials, where trials need to be added to the original design to compensate for the loss of statistical power; (v) display items, e.g., figures, that do not adequately show variability measures, scales, or topography; (vi) inconsistent use of terminology – particularly pertinent for ERPs and canonical MEEG frequency bands, which may change over the lifespan; (vii) underspecified or omitted results of statistical analyses – including model assumptions, test statistics, effect sizes, and statistical maps for mass-univariate and multivariate analyses (Šoški et al., 2021).

4. Beyond the signal

Each scientific discipline has its own distinctive “way of doing things”, and the more specialized a scientist becomes, the more they notice the devil in the details. Much more could be written about optimizing experimental designs, data collection, signal processing, or any other specific aspect of the MEEG research cycle, but this is beyond the general scope of this article. In fact, in this section, we take a further step back to focus on social and ethical aspects of our discipline in particular, and science more generally, because these aspects, too, should be addressed by GSP.

The general principle is one of caring: for our work, participants, societies, environments, and ourselves. As our branch of science deals directly with human beings, we have a duty of care for the duration that they are with us in the laboratory, and for as long any effects may linger. Our concern for human well-being however should extend beyond participants to all those potentially affected by our research, including our own selves and our wider social communities. This section will provide some perspectives on how a step back from scientifically focused details can improve our work, our field, our society, and more. Section 4.1 addresses how neuroscience may influence societies for good or for bad, as neurotechnology becomes increasingly capable and attracts increasing commercial interest. A critical, philosophical look on some aspects of progress in neuroscience is outlined in Section 4.2, emphasizing the need to question the very nature of the work, and ensure its relevance. In Section 4.3, *slow science* is discussed, a movement and way of thinking that aims not just to fight the symptoms of current issues, but to rethink academia in a broader light. Finally, Section 4.4 touches upon the ever-relevant, broader environmental considerations that need to be part of all decision making.

4.1. Social responsibilities and neuroethics

Hume (1739) famously argued for the distinction between “is” and “ought”, separating factual judgements from moral ones, and reinforcing the division of scientific disciplines in “pure” and “applied” sciences (Proctor, 1991; Kincaid et al., 2007). There was, and still persists, an idea that these “pure” sciences can be “value-free”, or isolated from any moral considerations (Douglas, 2014). Irrespective of this larger philosophical debate, GSP requires an awareness of the moral implications and potential (mis)uses of scientific

findings, even when no immediate application is being investigated. With direct-to-consumer neurotechnology becoming increasingly available to the general public (Ienca et al., 2018), and with large international social media and tech companies investing in neurotechnology (Moses et al., 2019; Musk, 2019), the possible ethical, legal, and societal implications of neuroscientific research are currently more pressing than ever.

BCIs, in particular, allow neural correlates of mental states to be identified and used in real time (Wolpaw and Wolpaw, 2012), and are a major focus of current consumer applications. With this technology, among other things, neuroprosthetics can be developed to support the motor-impaired (Wolpaw et al., 2002), and the devices we use every day can be made to automatically adapt to our mental states (Zander and Kothe, 2011). However, it has also been shown that information can be obtained from brain activity that people did not intend to communicate, or did not wish to reveal (Schultze-Kraft et al., 2016; Zander et al., 2016). As such, neurotechnology may present a danger to the privacy of thought (Mecacci and Haselager, 2019). Furthermore, it is possible for a BCI-based device to actively attempt to manipulate a human participant's mental state, or to extract specific information from their brain activity, potentially without their knowledge or consent (Fairclough, 2017; Krol et al., 2020). This approach is, for example, the basis of the "guilty knowledge test", which attempts to identify incriminating response patterns in a criminal suspect's evoked brain activity (Rosenfeld et al., 2008). It is not always clear whether such tests currently violate any existing rights, such as those that protect against unreasonable searches or self-incrimination (Pardo and Patterson, 2013). However, what is legal is not necessarily morally right, and regulatory law generally lags behind new scientific developments. In the meantime, we should thus consider whether or in what way our work might be used to violate not just existing laws, but also any generic humane principles or human rights that do not yet exist, but ought to exist. The right to *mental integrity* is one such proposed right that MEEG researchers may inadvertently violate, or allow others to violate (Ienca and Andorno, 2017). With this in mind, the Republic of Chile recently became the first country in the world to pass a specific "neurorights" law (Strickland and Gallucci, 2022). Other issues related to neuroscience concern the potential of "brain hacking" (Ienca and Haselager, 2016), the influence of neurotechnology on our sense of agency (Haselager, 2013), and the societal impact of cognitive enhancement (Hyman, 2011), to name a few.

As such, neither neuroscience in general, nor MEEG research in particular, exists in moral isolation: indeed, neuroethics has a long history of thought on this topic. Scientists and researchers are encouraged to identify and discuss the potential ethical, legal, and societal implications of their research in their publications. In this vein, the Organization for Economic Co-operation and Development (OECD) recently adopted the Recommendation on Responsible Innovation in Neurotechnology (OECD, 2019), which contains nine principles aimed at guiding researchers to minimize societal risks. Among other things, it asks researchers to consider the multidimensional societal implications that their work may have, to anticipate potential misuses, and to assess the safety of new neurotechnological developments. This process can be supported by, for example, the five criteria introduced by Mecacci and Haselager (2019) related to accuracy, reliability, informativity, concealability, and enforceability, suggesting that we answer such questions as: Can this technology we are developing be used without the knowledge of the subject? Can it be used against their will?

Here, the issues may not be obvious: even “classic” evoked potentials such as the P300 may be abused to extract information from the brain of unwitting participants (Martinovic et al., 2012) using methods such as cognitive probing (Krol et al., 2020).

Accepting that moral judgment of our work is unavoidable, we can proactively integrate ethical and societal considerations in our research and curricula, and look at neurotechnology not as a means (e.g., to gather data) but as an end: we can use the possibilities afforded by this technology to actively and explicitly “promote human flourishing” (Kellmeyer, 2018).

4.2. The map and the territory

Collecting clean data, computing accurate scores, and optimizing decision boundaries are some of the cornerstones of current scientific practice. However, results without a theory to interpret them are of limited interest, and there is currently a glaring need for robust psychological theories for progress in the cognitive sciences.

“... In that Empire, the Art of Cartography attained such Perfection that the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province. In time, those Unconscionable Maps no longer satisfied, and the Cartographers Guilds struck a Map of the Empire whose size was that of the Empire, and which coincided point for point with it. The following Generations, who were not so fond of the Study of Cartography as their Forebears had been, saw that that vast map was Useless, and not without some Pitilessness was it, that they delivered it up to the Inclemencies of Sun and Winters” (Borges, 1946)

This text of Borges was used by Umberto Eco for an essay *on the impossibility of drawing a map of the empire on a scale of 1 to 1* (Eco et al., 1995). There Eco argues that, once folded, the map may not depict the folded map itself and may thus be unfaithful and useless. This reminds us of Heisenberg’s uncertainty principle, in the sense that the potential accuracy of our knowledge is fundamentally limited. Eco goes on to state that when the map is installed it does not represent the map itself over the territory. This is a sort of Russell’s paradox applied to maps: a normal map cannot map itself. In other words: (a) every 1:1 map always reproduces the territory unfaithfully; (b) when the map is realized, the empire becomes unreproducible; and (c) every 1:1 map of the empire decrees the end of the empire as such and therefore is the map of a territory that is not an empire.

If we replace the word *empire* with the word *brain*, we come closer to the topic of interest here, and recall that we do not study the brain, but different representations of the brain. Paraphrasing Eco, recording all individual neurons of the brain, may decree the end of the scientific study of the brain as such and therefore a scientific field that is not concerned with the brain, but with something else that is qualitatively different from it (e.g., individual neurons), and that misses its higher order emerging properties (e.g., the mind).

This literary and philosophical escapade brings us to examine notions of *orders of magnitude* in science. In general one can observe an increase in the number of almost everything (but the questions): from single case (or single cell) studies to Big Data, with an increase in all numbers (of brains, of bytes, of CPUs), do we truly understand the brain –

not mentioning the mind – any better by just increasing numbers? Or can we learn to ask different scientific questions from Big Data, to solve deeper questions in a way we have not done previously?

These questions bring up the role of theory in scientific investigation. EEG and MEG as used in cognitive sciences rest on two fundamental theoretical domains. On the one hand, the MEEG physics is well understood and electro-magnetic fields are accurately predicted from the underlying neural sources (Ilmoniemi and Sarvas, 2019). Computational models with varying levels of detail therefore match observed MEEG fields with remarkable accuracy (see, e.g., Fig. 1 in Wendling et al. (2016)), or simulation frameworks such as the Human Neocortical Neurosolver, (Neymotin et al., 2019). On the other hand, cognitive MEEG studies are concerned with explaining psychological phenomena and, in this realm, build on much weaker and controversial theories. The “theory crisis” in psychology (Eronen and Bringmann, 2021) and beyond is limiting progress here, and arguably a fundamental cause of the replication crisis. A rigorous method for theory construction (e.g., Borsboom et al., 2021) is needed and should be part of our GSP, just like legends and annotations on maps make the art of cartography much more than merely capturing every detail of the world in an objective, but meaningless, way.

4.3. Rethinking our narratives: the challenge of “slow science”

Even as the open science movement participates in improving research quality, notably by developing GSP, the underlying causes of the need for such GSP in the first place remain largely unquestioned. In this section, some aspects of the *culture of speed*, the main research culture in academia, are described. As demonstrated with several examples, adopting the perspective of *slow science* can help challenge our research culture in a deeper way.

4.3.1. The negative consequences of the culture of speed and the limited response of the open science movement—Publication numbers in academia have exploded as more scientists are publishing more papers per capita than in the past (Bornmann and Mutz, 2015). Our current system, in which research is awarded in a way that favors this acceleration, shapes research practice dramatically. In particular, it favors a “culture of speed”: incentivizing data accumulation with no time for reflection, fostering a competition that obstructs collaboration, requiring time to write grant applications instead of actually doing science. It also has negative consequences for research quality: the use of metrics to decide grant funding or hiring encourages abuse (Smaldino and McElreath, 2016) and strategic gaming (Chapman et al., 2019), and decreases trust in publication quality (Vazire, 2017). Importantly, it also substantially impacts researchers’ mental health, with increased levels of stress and frustration as a consequence of lack of time, work-overload, often coupled with fixed-term contracts, in uncaring work environments (see Shaw and Ward 2014) for an overview of this topic).

Critically, this culture of speed creates strong incentives against the GSP presented in this document, which can lead to questionable research practices that have led to the replication crisis. The open science movement is the most widely accepted response to the replication crisis, and a strong remedy against questionable research practices. However,

by itself, it has a limited impact on these incentives. The incentives it offers to encourage openness, integrity, and reproducibility, and the tools it provides to solve current problems (reviewed in this paper and elsewhere, e.g., Renkewitz and Heene, 2019), may not suffice to address the roots of the culture of speed. Without a deeper reflection on these roots, newly created incentives may become tomorrow's new metrics, enabling "open science washing" or turning GSP into a list of boxes to tick to get one's work published or funded. As an example, the way in which the open science movement participates in the creation of new publication types (preprints, registered reports), contributing to the rise in publication numbers mentioned above, currently remains largely unquestioned. Despite offering positive steps towards more transparency and better scientific practice, it fails to address the origins of the problem.

4.3.2. What do we do? A slow science perspective—The slow science philosophy offers a broad response to the culture of speed (Stengers, 2013). Slow science aims to transform research and teaching into a "sustainable collective praxis" (Salo and Heikkinen, 2018). This movement emphasizes the collective aspects of science rather than protecting one's own ideas, fostering reflection about practices and the relevance of research questions. By definition, slow science promotes more open science, where the goal is not to try to 'fix' but to rethink academia.

Importantly, slow science is not about forcing scientists to slow down their publication pace, even though some explicitly suggested this might be a desirable option (Frith, 2020). It is also not about returning to some idealized golden age where scientists could focus deeply on their work, immune from the needs of society and productivity imperatives. Instead, slow science is about resisting the culture of speed and the idea of "wasting time" with futilities (Stengers, 2013). There is a need to question those narratives and how time pressure and time fragmentation (Ylijoki and Mäntylä, 2003) affects creativity and critical thinking on top of jeopardizing wellbeing (Maestre, 2019). For instance, overwork and multitasking are traditionally celebrated, and work-life balance and mental health have yet to be taken more seriously (Berg and Seeber, 2015). To foster those discussions and reflections, there have to be spaces where one can imagine and experiment with alternatives to the current system, even at a local level.

Initiatives already exist for paving the slow science way (see links in Appendix Table A1). Some examples are research frameworks that have emerged from group discussions at conferences: the San Francisco Declaration on Research Assessment (a.k.a. DORA) promotes ways of evaluating researchers that go beyond publication metrics, or the HIBAR Research Alliance, that aims to make research and innovation more integrated and aligned with society's critical problems. These initiatives have already been acknowledged by several institutions (see their respective lists of signatories or members). Other initiatives include discussion groups where scholars from different disciplines discuss and rethink academia and imagine a healthier research culture. Among them are such groups as Slow Science in Belgium or the Better Science initiative in Switzerland. Both groups have been closely working with their host universities to improve their working conditions or offer training programs. Finally, some researchers have created independent institutes: the Ronin Institute that promotes research outside traditional institutions, or IGDORE, the Institute

for Globally Distributed Open Research and Education, whose objective is to improve the quality of science, science education, and scientists' quality of life. Close to this philosophy, the Learning Planet Institute is exploring new ways of doing research, learning, or teaching while connecting with other parts of the society (industries, citizens, etc.). These institutes show how hybrid research, both inside and outside academia, is a viable option to consider, challenging the traditional pipeline in academia.

These existing networks, discussion groups, and more radical alternatives are all concrete examples that individual scientists can, at their level, promote healthier research practices and improve conditions. These examples should encourage scientists to explore other practices, question their relationship to academia, and start conversations with colleagues. For instance, the Wellcome Foundation developed a discussion kit, the Café Culture Initiative, to foster discussions about reimagining the research culture and working together.

In a nutshell, more needs to be done than just 'fixing academia' or improving scientific practice. Narratives, research frameworks, and practices need to be actively rethought. As summarized by Lancaster et al. (2018, p. 10), "*Making science better is not just about 'creating better incentives', but a collective cultural shift beyond viewing competition and individualistic success as the sole defining feature of science.*"

4.4. GSP for environmental sustainability

Finally, there remains an important aspect of GSP that we have not discussed. Current scientific practices, like many human activities, are on an unsustainable long-term path. The current ecological collapse (The Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES) report, Díaz et al., 2019) and climate change (Intergovernmental Panel on Climate Change (IPCC) reports) is leading planet Earth to a disastrous state. GSP for the future therefore includes immediate sustainable practices in all fields of science. To this end, many collective actions are taking place around the world. Over the past decade, most academic institutions and research foundations have created departments and initiatives dedicated to studying and acting for environmental sustainability on campuses and in research centers (see Appendix Table A1). These institutions, along with academic grassroots collective initiatives, are important local actors that could change GSP by setting local research agendas on more sustainable paths (Rae et al., 2022). As preliminary as these actions may be, they signal that the scientific community is progressively moving towards incorporating environmental considerations in the definition of GSP.

5. Discussion

This paper summarized select contributions to the 2020 LiveMEEG conference, which brought together international experts to discuss current progress and perspectives on good scientific practice in MEEG research. Virtual sessions and panels covered GSP across all stages of the research lifecycle, as well as broader topics beyond experimental work, such as societal responsibilities and research culture. Strikingly, a common theme emerged regarding the value of collaborative work. Many contributors emphasized the benefits to be had from reaching out beyond office, lab, and institutional walls, and beyond disciplines; from leveraging each other's competence, resources, and perspectives.

Indeed, there were many explicit suggestions of tools and methods for collaborative work. For example, collaborative meetings where scientists support each other to overcome human biases at early stages of experimental planning are proposed; pre-registration explicitly invites community feedback at early stages of the research process; different labs are encouraged to work together to reproduce findings and provide resources to the community; guides are being developed to cultivate further collaborative projects; and scientists across disciplines are reimagining science and academia in the current century.

Of course, not all work can be done collectively. Yet, even from those GSP perspectives that do not explicitly deal with shared work, a general theme of collaboration emerges – a common collaborative mindset, as it were, from which many different GSP guidelines emerge almost naturally. This mindset encourages scientists to see themselves not (primarily) as solitary experts in their specific field, but as part of a larger movement that ultimately shares the same goals of renewing truth and understanding. In essence, scientists are encouraged to either seek out collaborators where possible and appropriate, or otherwise imagine invisible or future collaborators who are eager to join or continue – but not usurp – the project.

The mere *idea* of having collaborators, imaginary or otherwise, almost automatically leads to practices that enable transparency, open science, and reproducibility. That is, assuming that the project at some point will be joined by or passed on to someone else provides clear, almost automatic incentives to uphold GSP: collaborators must be able to access previous work, understand the full experimental philosophy, retrace analysis steps, and verify outcomes. As such, all efforts made towards reproducibility, be it by following known protocols, by providing annotated data, legible code, or full open resources, are in essence collaborative endeavors, anticipating potential collaborators. Even scientific publication and citation practices can be seen as a continuous collaboration across time and space, building upon previous work and creating new work for others to continue to build upon. Finally, future collaborators should, wherever possible, be saved from the disappointment of finding an avoidable flaw in the work on which they are building their own, emphasizing the need for GSP also in the methodological details.

Now a key challenge for the shift towards a fully collaborative culture in MEEG is the need for coordination. Scientists have too long been trained to think and act individually, or in small mentor-advisee pairs, and generally lack training in cooperative action. Leading or contributing to a collective effort is in no way as easy and simple as conducting an individual project. This is perhaps where other areas of science and society can aid the MEEG community.

Thus, the collaborative mindset links perfectly with traditional GSP, but it also encourages scientists to “think big” in terms of developing new GSP, and guides collective efforts in open communities to develop worldwide standards, movements, and cultures. In summary, the many LiveMEEG 2020 contributors have shown that going forward, the search for “Eureka!”, or “I have found it”, is best envisioned as an endeavor to, instead, reach “we have found it” – “Eurékamen!”

Funding

GN was supported by Spanish Government through grant PID2020-119458RJ-I00 and by the AXA Research Fund; MAE was supported by UBC 4YF; SJL was supported by National Institute of Mental Health (R25MH080794 and R01MH087450); JCM was supported in part by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) of the National Institutes of Health (NIH) under award number R01EB026299; YGP was supported by Russian Foundation for Basic Research (RFBR) #1,901,300,027; AP, GN and MS are supported by an NIH CRCNS: US-France Data Sharing Proposal (NIBIB (USA) R01 EB030896 and ANR-20-NEUC-0004-01); AS is employed as Coordinator Open Science at Erasmus Research Services (Erasmus University Rotterdam); WSA was supported by the John Templeton Foundation; BS was supported by The Artificial and Natural Intelligence Toulouse Institute; RT was supported by British Academy Postdoctoral Fellowship (SUAI/028 RG94188 pf170046); MC was supported by Program Investissements d'avenir ANR - 10 - IAIHU - 06; CF was supported by an ANR-20-CE28-0017.

Appendix. Table A1

Table A1

Example resources for Good Scientific Practice in MEEG research.

This article is based on the contributions of the LiveMEEG 2020 online conference		
<i>LiveMEEG conference</i>	livemeeg2020.org	Program and recorded sessions
Early considerations for MEEG projects		
Human Factors		
<i>Wikipedia's list of cognitive biases</i>	en.wikipedia.org/wiki/List_of_cognitive_biases	<i>A collaborative listing with over 100 documented (not all strictly distinct) cognitive biases</i>
<i>The cognitive bias cheat sheet</i>	betterhumans.pub/cognitive-bias-cheat-sheet-55a472476b18	<i>A highly viewed (non academic) blog post structuring all cognitive biases of Wikipedia's list of cognitive biases, with an appealing visualization.</i>
<i>Coursera MOOC by Sinnott-Armstrong</i>	coursera.org/instructor/~932346	<i>An in depth practical introduction to arguments, proper reasoning and logical thinking</i>
<i>Rain Cloud Plots</i>	github.com/RainCloudPlots/RainCloudPlots	<i>A a multi-platform tool for robust data visualization</i>
Preregistration		
<i>OSF registries</i>	osf.io/registries	<i>Pre-registration platform of OSF</i>
<i>Clinical trials</i>	clinicaltrials.gov	<i>Pre-registration platform</i>
<i>As Predicted</i>	aspredicted.org	<i>Pre-registration platform</i>
<i>Guidelines for ERP preregistration</i>	10.1016/j.ijpsycho.2021.02.016	<i>Guidelines for ERP preregistration</i>
Reproducibility		
<i>The Turing Way</i>	the-turing-way.netlify.app/welcome	<i>Open source community-driven guide to reproducible, ethical, inclusive and collaborative data science</i>
<i>ERP CORE</i>	erpinfo.org/erp-core	<i>Compendium of Open Resources and Experiments</i>
<i>#EEGManyLabs</i>	osf.io/yb3pq	<i>A large scale crowdsourced replication project</i>
<i>EEGManyPipelines</i>	eegmanypipelines.org	<i>Project involving many independent analysis teams to investigate how analysis approaches affect results</i>
<i>MOABB (Mother of all BCI Benchmarks)</i>	github.com/NeuroTechX/moabb	<i>Platform to build a comprehensive benchmark of popular BCI algorithms</i>

		<i>applied on an extensive list of freely available EEG datasets</i>
<i>BNA Credibility</i>	bnacredibility.org.uk	<i>A community working together to increase the credibility of research</i>
General purpose platforms for open research		
<i>Open Science Framework (OSF)</i>	osf.io	<i>A popular, general purpose platform for collaborative project tracking, organization, storage, and referencing</i>
<i>Center for Open Science</i>	cos.io	<i>A non profit organization committed to increase openness, integrity, and reproducibility of research.</i>
<i>Zenodo</i>	zenodo.org	<i>An open data platform supported by the European Commission to advance their open data policy</i>
<i>FigShare</i>	figshare.com	<i>Platform to upload any file format and make research outputs citable, shareable and discoverable</i>
<i>GitHub</i>	github.com	<i>is a provider of Internet hosting for software development and version control using Git.</i>
<i>GitLab</i>	gitlab.com	<i>is a web-based DevOps lifecycle tool that provides a Git repository manager providing wiki, issue-tracking and continuous integration and deployment</i>
Collecting, analyzing, reporting, and sharing		
<hr/>		
Data Collection		
<i>EEG recording protocol</i>	protocolexchange.researchsquare.com/article/pex-974/v2	<i>Protocol for Reducing COVID-19 Transmission Risk in EEG Research</i>
<i>LSL</i>	labstreaminglayer.readthedocs.io	<i>Lab Streaming Layer (LSL) is a system for the unified collection of measurement time series in research experiments</i>
Stimulus presentation		
<i>Psychophysics toolbox</i>	psyctoolbox.org	<i>A free set of Matlab and GNU Octave functions for stimulus presentation in vision and neuroscience research.</i>
<i>PsychoPy</i>	psycho.py	<i>Free cross-platform package allowing to run a wide range of experiments in the behavioral sciences.</i>
<i>Pavlovia</i>	pavlovia.org	<i>Platform to run, share, and explore behavioural experiments online</i>
Simulations		
<i>EEGSourceSim</i>	osf.io/fmuae	<i>MATLAB-based toolbox providing individual head models and corresponding atlases to simulate noise, SSEP, and multivariate autoregressive signals.</i>
<i>SEREEGA</i>	github.com/lrkrol/SEREEGA	<i>MATLAB-based general multipurpose toolbox to simulate event-related EEG activity, also available as EEGLAB plugin.</i>
<i>simBCI</i>	gitlab.inria.fr/sb/simbci	<i>MATLAB-based toolbox focusing on testing BCI classification methods.</i>
<i>SimEEG</i>	audiospeech.ubc.ca/research/brane/brane-lab-software	<i>MATLAB-based toolbox with a focus on simulating interacting brain signals. Now integrated with Brainstorm</i>
<i>The Virtual Brain</i>	thevirtualbrain.org	<i>Python-based platform to simulate whole-brain dynamics using structurally connected network models.</i>

Data Sharing		
<i>Open Brain Consent</i>	open-brain-consent.readthedocs.io/en/stable	<i>Sample consent forms to facilitate data sharing in neuroimaging which have been previously approved by ethic committees in different institutions</i>
<i>BIDS</i>	bids-standard.org	<i>The Brain Imaging Data Structure</i>
MEEG data repositories		
<i>HCP-YA</i>	humanconnectome.org/study/hcp-young-adult	<i>Human Connectome Project healthy young adults</i>
<i>CamCAN</i>	cam-can.org	<i>The Cambridge center for Ageing and Neuroscience (Cam-CAN). It contains epidemiological cognitive, and neuroimaging including rest and task resting-state MEG data</i>
<i>Donders Repository</i>	data.donders.ru.nl	<i>An institutional data sharing repository with more than 150 published neuroimaging MEG and EEG datasets</i>
<i>OpenNeuro</i>	openneuro.org	<i>A free and open platform for sharing MRI, MEG, EEG, iEEG, ECoG, and ASL data following BIDS</i>
<i>OMEGA</i>	mcgill.ca/bic/resources/omega	<i>The Open MEG Archive. A continuously expanding repository that contains MEG, EEG, anatomical MRI volumes, demographic and questionnaire information on hundreds of participants</i>
<i>BioFIND</i>	doi.org/10.1101/2021.05.19.21257330	<i>A multi-site, multi-participant magnetoencephalography resting-state dataset to study dementia</i>
<i>CHBMP</i>	doi.org/10.1101/2020.07.08.194290	<i>Open multimodal neuroimaging and cognitive dataset from 282 healthy participants</i>
<i>MOUS</i>	doi.org/10.1038/s41597-019-0020-y	<i>A 204-subject multimodal neuroimaging dataset to study language processing (MRI+MEG)</i>
<i>Temple University EEG corpus</i>	isip.piconepress.com/projects/tuh_eeg/html/downloads.shtml	<i>Electroencephalography (EEG) Resources</i>
<i>List of open datasets in ephys</i>	github.com/openlists/ElectrophysiologyData	<i>List of openly available electrophysiological data, including EEG, MEG, ECoG/iEEG, and LFP data (by Tom Donoghue)</i>
Software		
<i>Brainstorm</i>	neuroimage.usc.edu/brainstorm	<i>Brainstorm main page</i>
<i>EEGLAB</i>	sccn.ucsd.edu/eeglab	<i>EEGLAB main page</i>
<i>FieldTrip</i>	fieldtriptoolbox.org/	<i>FieldTrip</i>
<i>MNE-python</i>	mne.tools	<i>MNE-python</i>
<i>SPM</i>	fil.ion.ucl.ac.uk/spm	<i>SPM</i>
<i>MMVT</i>	mmvt.mgh.harvard.edu	<i>visualization tool for multimodal imaging</i>
<i>MNELAB</i>	pypi.org/project/mnelab	<i>MNELAB is a graphical user interface for MNE</i>
Cloud computing platforms		
<i>brainlife</i>	brainlife.io	<i>Free and open-source cloud platform for secure neuroscience data analysis.</i>
<i>Neuroscience Gateway (NSG)</i>	nsgportal.org	<i>High Performance Computing (HPC) for neuroscience researchers</i>

<i>The Canadian Open Neuroscience Platform (CONP)</i>	conp.ca	<i>The Canadian Open Neuroscience Platform (CONP)</i>
Statistics		
<i>Standardized Measurement Error</i>	doi.org/10.18115/D58G91	<i>Standardized Measurement Error (SME) demo scripts</i>
<i>Super Power</i>	aaroncaldwell.us/Superpower	<i>R package to simulate factorial designs and empirically calculate power</i>
<i>Super Power Book</i>	aaroncaldwell.us/SuperpowerBook	<i>Book describing the capabilities of the SuperPower R package</i>
<i>simstudy</i>	cran.r-project.org/web/packages/simstudy	<i>Simulation of Study Data</i>
<i>simr</i>	github.com/pitakakariki/simr	<i>Power Analysis of Generalised Linear Mixed Models by Simulation</i>
<i>Nilearn</i>	nilearn.github.io	<i>Tool for Statistics for NeuroImaging in Python</i>
<i>Scikit-Learn</i>	scikit-learn.org	<i>Tool for Machine Learning in Python</i>
Standards		
<i>BIDS</i>	bids-standard.org	<i>The Brain Imaging Data Structure (BIDS), for describing and organizing neuroimaging data</i>
<i>BIDS Specification</i>	bids-specification.readthedocs.io	<i>It included BIDS core specification as well as many modality-specific extensions</i>
<i>BIDS Starter Kit</i>	github.com/bids-standard/bids-starter-kit	<i>A community-curated collection of tutorials, wikis, and templates to get started with BIDS</i>
<i>MNE-BIDS</i>	mne.tools/mne-bids/stable/index.html	<i>BIDS converter for MEEG data as part of the MNE-Python software</i>
<i>data2bids</i>	fieldtriptoolbox.org/reference/data2bids	<i>BIDS converter for MEEG data as part of the FieldTrip toolbox</i>
<i>bids-matlab-tools</i>	github.com/sccn/bids-matlab-tools	<i>BIDS converter for EEG data as part of the EEGLAB toolbox</i>
<i>Brainstorm BIDS tools</i>	neuroimage.usc.edu/brainstorm/ExportBids	<i>BIDS converter for MEEG data as part of the Brainstorm toolbox</i>
<i>BIDS examples</i>	github.com/bids-standard/bids-examples	<i>A set of BIDS-compatible datasets with empty raw data files</i>
<i>pybids</i>	github.com/bids-standard/pybids	<i>Python library to centralize interactions with datasets conforming BIDS</i>
<i>bids-matlab</i>	github.com/bids-standard/bids-matlab	<i>This repository aims at centralising MATLAB/Octave tools to interact with datasets conforming to BIDS</i>
<i>BIDS Apps</i>	bids-apps.neuroimaging.io	<i>Portable neuroimaging pipelines that understand BIDS datasets</i>
<i>COBIDAS MEEG</i>	cobidasmeeg.wordpress.com	<i>Blog describing the COBIDAS MEEG best practices</i>
<i>eCOBIDAS</i>	remi-gau.github.io/eCobidas	<i>electronic Python-based user-friendly checklist for COBIDAS</i>
<i>ARTEM-IS</i>	osf.io/w4nt6	<i>Template designed to supplement a paper or a preregistration and make reporting EEG methodology easier and more accurate by providing specific fields for specific details</i>
Preprint servers		
<i>OSF preprints</i>	osf.io/preprints	<i>Preprint server in OSF</i>
<i>bioRxiv</i>	biorxiv.org	<i>Preprint server for biology</i>
<i>medRxiv</i>	medrxiv.org	<i>Preprint server for health sciences</i>

<i>PsyArXiv</i>	psyarxiv.com	<i>Preprint server for the psychological sciences</i>
<i>arXiv</i>	arxiv.org	<i>Open-access archive for scholarly articles in multiple fields (e.g. physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering and systems science, and economics)</i>
<i>NeuroLibre</i>	neurolibre.org	<i>Preprint server for interactive data analyses</i>
Beyond the signal		
<i>Slow Science Perspective</i>		
<i>DORA initiative</i>	sfedora.org	<i>The Declaration on Research Assessment (DORA) recognizes the need to improve the ways in which researchers and the outputs of scholarly research are evaluated</i>
<i>HIBAR Research Alliance</i>	hibar-research.org	<i>Highly Integrative Basic and Responsive (HIBAR) research to contribute to solving society's critical problems</i>
<i>Ronin institute</i>	ronininstitute.org	<i>New model for scholarly research recognizes the value of people outside of traditional academia</i>
<i>Slow Science in Belgium</i>	slowscience.be	<i>An interuniversity platform for discussion on academia's future</i>
<i>Better Science</i>	betterscience.ch	<i>Contributes to a rethinking of the current paradigm of quantifiable scientific and scholarly work</i>
<i>IGDORE</i>	igdore.org	<i>Institute for Globally Distributed Open Research and Education (IGDORE) aims to improve the quality of science, science education, and quality of life for scientists, students and families</i>
<i>Learning Planet Institute</i>	learningplanetinstitute.org	<i>Promotes learning, research, collective intelligence, and creativity to help with the complex challenges of the changing world, as a learning-society revolution</i>
<i>Café Culture</i>	wellcome.org/what-we-do/our-work/research-culture/hosting-your-cafe-culture-discussion	<i>Platform to discuss about the challenges in research culture, reflect on a better culture and solutions for change</i>
<i>Ecological responsibility</i>		
<i>Sustainability network of the Max Planck Society</i>	nachhaltigkeitsnetzwerk.mpg.de	<i>A leading network on sustainability studies</i>
<i>Tyndall</i>	tyndall.ac.uk	<i>A major european research center on climate change</i>
<i>Labos1point5</i>	labos1point5.org	<i>Research initiative to reduce the Green House Gas emissions of research in the public sector</i>
<i>OHBM environment action group</i>	ohbm-environment.org	<i>OHBM's initiative to reduce the impact of neuroimaging research on the environment</i>
<i>NoFlyClimateSci</i>	noflyclimatesci.org	<i>Grassroots academic initiative of scientists who decided to fly less and push for climate initiatives through their institutions</i>
<i>Scientists4Future</i>	scientists4future.org	<i>International network of scientists supporting the global climate movement by providing facts and materials based on reliable and accepted scientific data</i>

References

- Abelson RP, 1995. *Statistics as Principled Argument*, 1st ed. Psychology Press doi: 10.4324/9781410601155.
- Allen C, Mehler DMA, 2019. Open science challenges, benefits and tips in early career and beyond. *PLoS Biol.* 17 (5), e3000246. doi: 10.1371/journal.pbio.3000246. [PubMed: 31042704]
- Allen M, Poggiali D, Whitaker K, Marshall TR, Langen JV, Kievit RA, 2021. Rain-cloud plots: a multi-platform tool for robust data visualization (4:63). *Wellcome Open Res.* doi: 10.12688/wellcomeopenres.15191.2.
- Ambekar A, Ward C, Mohammed J, Male S, Skiena S, 2009. Name-ethnicity classification from open sources. In: *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 49–58. doi: 10.1145/1557019.1557032.
- Andersen LM, 2018a. Group analysis in MNE-python of evoked responses from a tactile stimulation paradigm: a pipeline for reproducibility at every step of processing, going from individual sensor space representations to an across-group source space representation. *Front. Neurosci.* 12, 6. doi: 10.3389/fnins.2018.00006. [PubMed: 29403349]
- Andersen LM, 2018b. Group analysis in FieldTrip of time-frequency responses: a pipeline for reproducibility at every step of processing, going from individual sensor space representations to an across-group source space representation. *Front. Neurosci.* 12, 261. [PubMed: 29765297]
- Appelhoff S, Sanderson M, Brooks T, van Vliet M, Quentin R, Holdgraf C, Chaumon M, Mikulan E, Tavabi K, Höchenberger R, Welke D, Brunner C, Rockhill A, Larson E, Gramfort A, Jas M, 2019. MNE-BIDS: organizing electrophysiological data into the BIDS format and facilitating their analysis. *J. Open Source Softw.* 4 (44), 1896. doi: 10.21105/joss.01896. [PubMed: 35990374]
- Bae GY, Olkkonen M, Allred SR, Wilson C, Flombaum JI, 2014. Stimulus-specific variability in color working memory with delayed estimation. *J. Vis* 14 (4). doi: 10.1167/14.4.7, 7–7.
- Bagic AI, Knowlton RC, Rose DF, Ebersole JS, 2011. American clinical magnetoencephalography society clinical practice guideline 1: recording and analysis of spontaneous cerebral activity. *J. Clin. Neurophysiol.* 28 (4), 348–354. doi: 10.1097/WNP.0b013e31822727fe. [PubMed: 21811121]
- Baillet S, Moshier J, Leahy R, 2000. BrainStorm beta release: a Matlab software package for MEG signal processing and source localization and visualization. *Neuroimage* 11. doi: 10.1016/S1053-8119(00)91843-3.
- Baker DH, Vilidaite G, Lygo FA, Smith AK, Flack TR, Gouws AD, Andrews TJ, 2020. Power contours: optimising sample size and precision in experimental psychology and human neuroscience. *Psychol. Methods* doi: 10.1037/met0000337.
- Bennett CM, Wolford GL, Miller MB, 2009. The principled control of false positives in neuroimaging. *Soc. Cogn. Affect. Neurosci.* 4 (4), 417–422. doi: 10.1093/scan/nsp053. [PubMed: 20042432]
- Baron J, 2006. *Thinking and Deciding*. Cambridge University Press.
- Berg M, Seeber BK, 2015. *The Slow professor: Challenging the Culture of Speed in the Academy*. University of Toronto Press doi: 10.3138/9781442663091.
- Bertolero MA, Dworkin JD, David SU, Lloreda CL, Srivastava P, Stiso J, Zhou D, Dzirasa K, Fair DA, Kaczkurkin AN, Marlin BJ, Shohamy D, Uddin LQ, Zurn P, Bassett DS, 2020. Racial and Ethnic Imbalance in Neuroscience Reference Lists and Intersections With Gender. *bioRxiv* (p. 2020.10.12.336230) doi: 10.1101/2020.10.12.336230.
- Birhane A, 2021. Algorithmic injustice: a relational ethics approach. *Patterns* 2 (2), 100205. doi: 10.1016/j.patter.2021.100205. [PubMed: 33659914]
- Bolt T, Nomi JS, Bzdok D, Uddin LQ, 2021. Educating the future generation of researchers: a cross-disciplinary survey of trends in analysis methods. *PLoS Biol.* 19 (7), e3001313. doi: 10.1371/journal.pbio.3001313. [PubMed: 34324488]
- Borges JL, 1946. *Del rigor en la ciencia*. Los Anales de Buenos Aires.
- Borland D, Taylor II RM, 2007. Rainbow color map (still) considered harmful. *IEEE Comput. Graph. Appl.* 27 (2), 14–17. doi: 10.1109/MCG.2007.323435.
- Bornmann L, Mutz R, 2015. Growth rates of modern science: a bibliometric analysis based on the number of publications and cited references. *J. Assoc. Inf. Sci. Technol.* 66 (11), 2215–2222. doi: 10.1002/asi.23329.

- Borsboom D, van der Maas HLJ, Dalege J, Kievit RA, Haig BD, 2021. Theory construction methodology: a practical framework for building theories in psychology. *Perspect. Psychol. Sci.* 16 (4), 756–766. doi: 10.1177/1745691620969647. [PubMed: 33593167]
- Boudewyn MA, Luck SJ, Farrens JL, Kappenman ES, 2018. How many trials does it take to get a significant ERP effect? It depends. *Psychophysiology* 55 (6), e13049. doi: 10.1111/psyp.13049. [PubMed: 29266241]
- Bridges D, Pitiot A, MacAskill MR, Peirce JW, 2020. The timing mega-study: comparing a range of experiment generators, both lab-based and online. *PeerJ* 8, e9414. doi: 10.7717/peerj.9414. [PubMed: 33005482]
- Brouwers K, Cooke A, Chambers CD, Henson R, Tibon R, 2020. Evidence for prereg posters as a platform for preregistration. *Nat. Hum. Behav.* 4 (9), 884–886. doi: 10.1038/s41562-020-0868-z. [PubMed: 32341494]
- Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR, 2013. Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14 (5), 365–376. doi: 10.1038/nrn3475. [PubMed: 23571845]
- Buzsáki G, Anastassiou CA, Koch C, 2012. The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nat. Rev. Neurosci.* 13 (6), 407–420. doi: 10.1038/nrn3241. [PubMed: 22595786]
- Casadevall A, Steen RG, Fang FC, 2014. Sources of error in the retracted scientific literature. *FASEB J.* 28 (9), 3847–3855. doi: 10.1096/fj.14-256735. [PubMed: 24928194]
- Chambers CD, 2013. Registered reports: a new publishing initiative at cortex. *Cortex* 49 (3), 609–610. doi: 10.1016/j.cortex.2012.12.016. [PubMed: 23347556]
- Chambers JM, Cleveland WS, Kleiner B, Tukey PA, 2018. *Graphical Methods for Data Analysis*. CRC Press doi: 10.1201/9781351072304.
- Chambers CD, Tzavella L, 2022. The past, present and future of Registered Reports. *Nat Hum Behav* 6, 29–42. doi: 10.1038/s41562-021-01193-7. [PubMed: 34782730]
- Chapman CA, Bicca-Marques JC, Calvignac-Spencer S, Fan P, Fashing PJ, Gogarten J, Guo S, Hemingway CA, Leendertz F, Li B, Matsuda I, Hou R, Serio-Silva JC, Chr. Stenseth N, 2019. Games academics play and their consequences: how authorship, h-index and journal impact factors are shaping the future of academia. *Proc. R. Soc. B Biol. Sci.* 286 (1916), 20192047. doi: 10.1098/rspb.2019.2047.
- Chatterjee P, Werner RM, 2021. Gender disparity in citations in high-impact journal articles. *JAMA Netw. Open* 4 (7), e2114509. doi: 10.1001/jamanet-workopen.2021.14509. [PubMed: 34213560]
- Chaumon M, Puce A, George N, 2021. Statistical power: implications for planning MEG studies. *Neuroimage* 233, 117894. doi: 10.1016/j.neuroimage.2021.117894. [PubMed: 33737245]
- Choy T, Baker E, Stavropoulos K, 2021. Systemic racism in EEG research: considerations and potential solutions. *Affect. Sci.* doi: 10.1007/s42761-021-00050-0.
- Cichy RM, Pantazis D, Oliva A, 2014. Resolving human object recognition in space and time. *Nat. Neurosci.* 17 (3), 455–462. doi: 10.1038/nn.3635. [PubMed: 24464044]
- Clayson PE, Baldwin SA, Larson MJ, 2021a. The open access advantage for studies of human electrophysiology: impact on citations and Altmetrics. *Int. J. Psychophysiol.* 164, 103–111. doi: 10.1016/j.ijpsycho.2021.03.006. [PubMed: 33774077]
- Clayson PE, Baldwin SA, Rocha HA, Larson MJ, 2021b. The data-processing multiverse of event-related potentials (ERPs): a roadmap for the optimization and standardization of ERP processing and reduction pipelines. *Neuroimage* 245, 118712. doi: 10.1016/j.neuroimage.2021.118712. [PubMed: 34800661]
- Clayson PE, Carbine KA, Baldwin SA, Larson MJ, 2019. Methodological reporting behavior, sample sizes, and statistical power in studies of event-related potentials: barriers to reproducibility and replicability. *Psychophysiology* 56 (11). doi: 10.1111/psyp.13437.
- Clayson PE, Keil A, Larson MJ, 2022. Open science in human electrophysiology. *Int. J. Psychophysiol.* 174, 43–46. doi: 10.1016/j.ijpsycho.2022.02.002. [PubMed: 35150771]
- Cohen MX, 2014. *Analyzing Neural Time Series Data: Theory and Practice*. MIT Press doi: 10.7551/mitpress/9609.001.0001.

- Combrisson E, Jerbi K, 2015. Exceeding chance level by chance: the caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *J. Neurosci. Methods* 250, 126–136. doi: 10.1016/j.neumeth.2015.01.010. [PubMed: 25596422]
- Community TTW, Arnold B, Bowler L, Gibson S, Herterich P, Higman R, Krystalli A, Morley A, O'Reilly M, Whitaker K, 2019. *The Turing Way: A Handbook for Reproducible Data Science*. Zenodo doi: 10.5281/zenodo.3233986.
- Cooper PS, Baillet S, Maroun REK, & Chong TT (2021). Over the rainbow: Guidelines for meaningful use of colour maps in neurophysiology. *NeuroImage*, 245, 118628. 10.1016/j.neuroimage.2021.118628 [PubMed: 34637902]
- Cramer AOJ, van Ravenzwaaij D, Matzke D, Steingroever H, Wetzels R, Grasman RPPP, Waldorp LJ, Wagenmakers EJ, 2016. Hidden multiplicity in exploratory multiway ANOVA: prevalence and remedies. *Psychon. Bull. Rev.* 23 (2), 640–647. doi: 10.3758/s13423-015-0913-5. [PubMed: 26374437]
- Dalal SS, Baillet S, Adam C, Ducorps A, Schwartz D, Jerbi K, Bertrand O, Garnero L, Martinerie J, Lachaux J-P, 2009. Simultaneous MEG and intracranial EEG recordings during attentive reading. *Neuroimage* 45 (4), 1289–1304. doi: 10.1016/j.neuroimage.2009.01.017. [PubMed: 19349241]
- Dale AM, Fischl B, Sereno MI, 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9 (2), 179–194. doi: 10.1006/nimg.1998.0395. [PubMed: 9931268]
- Dale AM, Sereno MI, 1993. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. *J. Cogn. Neurosci.* 5 (2), 162–176. doi: 10.1162/jocn.1993.5.2.162. [PubMed: 23972151]
- Delorme A, Makeig S, 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134 (1), 9–21. doi: 10.1016/j.jneumeth.2003.10.009. [PubMed: 15102499]
- Delorme A, Mullen T, Kothe C, Acar ZA, Bigdely-Shamlo N, Vankov A, Makeig S, 2011. EEGLAB, SIFT, NFT, BCILAB, and ERICA: new tools for advanced EEG processing. *Intell. Neurosci* 2011 (10). doi: 10.1155/2011/130714, 10–1010.
- Delorme A, Truong D, Martinez-Cancino R, Pernet C, Sivagnanam S, Yoshimoto K, Poldrack R, Majumdar A, Makeig S, 2021. Tools for importing and evaluating BIDS-EEG formatted data. In: *Proceedings of the 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 210–213. doi: 10.1109/NER49283.2021.9441399.
- Delorme A, Oostenveld R, Tadel F, Gramfort A, Nagarajan S, & Litvak V (2022). Editorial: From Raw MEG/EEG to Publication: How to Perform MEG/EEG Group Analysis With Free Academic Software. *Frontiers in Neuroscience*, 16. doi: 10.3389/fnins.2022.854471.
- Devezer B, Navarro DJ, Vandekerckhove J, Ozge Buzbas E, 2020. The case for formal methodology in scientific reform. *R. Soc. Open Sci.* 8 (3), 200805. doi: 10.1098/rsos.200805.
- Díaz S, Settele J, Brondízio ES, Ngo HT, Agard J, Arneth A, Balvanera P, Brauman KA, Butchart SHM, Chan KMA, Garibaldi LA, Ichii K, Liu J, Subramanian SM, Midgley GF, Miloslavich P, Molnár Z, Obura D, Pfaff A, ... Zayas CN, 2019. Pervasive human-driven decline of life on earth points to the need for transformative change. *Science* 366 (6471). doi: 10.1126/science.aax3100.
- Donchin E, Callaway E, Cooper R, Desmedt JE, Goff WR, Hillyard SA, Sutton S, 1977. Publication criteria for studies of evoked potentials (EP) in man: methodology and publication criteria. In: *Progress in Clinical neurophysiology: Vol. 1. Attention, Voluntary Contraction and Event-Related Cerebral Potentials*. Karger, pp. 1–11.
- Donoghue T, Haller M, Peterson EJ, Varma P, Sebastian P, Gao R, Noto T, Lara AH, Wallis JD, Knight RT, Shestyuk A, Voytek B, 2020. Parameterizing neural power spectra into periodic and aperiodic components. *Nat. Neurosci.* 23 (12), 1655–1665. doi: 10.1038/s41593-020-00744-x. [PubMed: 33230329]
- Douglas T, 2014. The dual-use problem, scientific isolationism and the division of moral labour. *Monash Bioethics Review* 32 (1–2), 86–105. doi: 10.1007/s40592-014-0004-9. [PubMed: 25434066]
- Dubarry AS, Badier JM, Trébuchon-Da Fonseca A, Gavaret M, Carron R, Bartolomei F, Liégeois-Chauvel C, Régis J, Chauvel P, Alario FX, Bénar CG, 2014. Simultaneous recording of MEG,

- EEG and intracerebral EEG during visual stimulation: from feasibility to single-trial analysis. *Neuroimage* 99, 548–558. doi: 10.1016/j.neuroimage.2014.05.055. [PubMed: 24862073]
- Dubarry AS, Liégeois-Chauvel C, Trébuchon A, Bénar C, & Alario FX (2022). An open-source toolbox for Multi-patient Intracranial EEG Analysis (MIA). *NeuroImage*, 119251. doi: 10.1016/j.neuroimage.2022.119251. [PubMed: 35568349]
- Duncan CC, Barry RJ, Connolly JF, Fischer C, Michie PT, Näätänen R, Polich J, Reinvang I, Van Petten C, 2009. Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clin. Neurophysiol.* 120 (11), 1883–1908. doi: 10.1016/j.clinph.2009.07.045. [PubMed: 19796989]
- Dworkin J, Zurn P, Bassett DS, 2020. (In)citing action to realize an equitable future. *Neuron* 106 (6), 890–894. doi: 10.1016/j.neuron.2020.05.011. [PubMed: 32553205]
- Eco U, Sterling D, Weaver W, 1995. *How to Travel with a Salmon & Other Essays*, 1st ed. Mariner Books.
- Eronen MI, Bringmann LF, 2021. The theory crisis in psychology: how to move forward. *Perspect. Psychol. Sci.* 16 (4), 779–788. doi: 10.1177/1745691620970586. [PubMed: 33513314]
- Es MWJV, Spaak E, Schoffelen JM, Oostenveld R, 2021. Reducing the efforts to create reproducible analysis code with FieldTrip. *BioRxiv* doi: 10.1101/2021.02.05.429886, 2021.02.05.429886.
- Etienne A, Laroia T, Weigle H, Afelin A, Kelly SK, Krishnan A, Grover P, 2020. Novel Electrodes for Reliable EEG Recordings on Coarse and Curly Hair. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6151–6154. doi: 10.1109/EMBC44109.2020.9176067 IEEE Engineering in Medicine and Biology Society. Annual International Conference, 2020.
- Fahimi Hnazaee M, Wittevrongel B, Khachatryan E, Libert A, Carrette E, Dauwe I, Meurs A, Boon P, Van Roost D, Van Hulle MM, 2020. Localization of deep brain activity with scalp and subdural EEG. *Neuroimage* 223, 117344. doi: 10.1016/j.neuroimage.2020.117344. [PubMed: 32898677]
- Fairclough SH, Jeon M, 2017. Physiological computing and intelligent adaptation. In: *Emotions and Affect in Human Factors and Human-Computer Interaction*. Academic Press, pp. 539–556. doi: 10.1016/B978-0-12-801851-4.00020-3.
- Fischl B, 2012. FreeSurfer. *Neuroimage* 62 (2), 774–781. doi: 10.1016/j.neuroimage.2012.01.021. [PubMed: 22248573]
- Franconeri SL, Padilla LM, Shah P, Zacks JM, Hullman J, 2021. The science of visual data communication: what works. *Psychol. Sci. Public Interest* doi: 10.1177/15291006211051956.
- Friston KJ, Ashburner J, Kiebel SJ, Nichols TE, Penny W, 2007. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press.
- Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak R, 1994. Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2 (4), 189–210. doi: 10.1002/hbm.460020402.
- Frith U, 2020. Fast lane to slow science. *Trends Cogn. Sci.* 24 (1), 1–2. doi: 10.1016/j.tics.2019.10.007. [PubMed: 31744772]
- Fulvio JM, Akinola I, Postle BR, 2021. Gender (Im)balance in citation practices in cognitive neuroscience. *J. Cogn. Neurosci.* 33 (1), 3–7. doi: 10.1162/jocn_a_01643. [PubMed: 33078992]
- Garrett-Ruffin S, Hindash AC, Kaczurkin AN, Mears RP, Morales S, Paul K, Pavlov YG, Keil A, 2021. Open science in psychophysiology: an overview of challenges and emerging solutions. *Int. J. Psychophysiol.* 162, 69–78. doi: 10.1016/j.ijpsycho.2021.02.005. [PubMed: 33556468]
- Farrens JL, Simmons AM, Luck SJ, Kappenman ES, 2020. Electroencephalogram (EEG) recording protocol for cognitive and affective human neuroscience research. *Protocol Exchange* doi: 10.21203/rs.2.18328/v3.
- Gau R, Gould van Praag C, van Mourik T, Wiebels K, Adolff F, Scarpazza C, Ruotsalainen I, Tepper A, Sjoerds Z, Simon J, Klapwijk E, Hortensius R, Bartlett J, & Moreau D 2019. COBIDAS checklist. doi: 10.17605/OSF.IO/ANVQY.
- Gau Rémi, Flandin Guillaume, Janke Andrew, tanguyduval Oostenveld, Robert Madan, Christopher Niso, Guiomar Szczepanik, Michał Mutsaerts, Henk Belyi, Nikita Norgaard, Martin Pernet, Cyril, & Chrisophe, Phillips. (2022). *bidsmatlab* (0.1.0). Zenodo. doi: 10.5281/zenodo.5910585.

- Gavaret M, Dubarry AS, Carron R, Bartolomei F, Trébuchon A, Bénar CG, 2016. Simultaneous SEEG-MEG-EEG recordings overcome the SEEG limited spatial sampling. *Epilepsy Res.* 128, 68–72. doi: 10.1016/j.eplepsyres.2016.10.013. [PubMed: 27816896]
- Gelman A, Loken E, 2013. The Garden of Forking paths: Why multiple Comparisons Can Be a problem, Even When There is No “fishing expedition” Or “p-hacking” and the Research Hypothesis Was Posited Ahead of Time. Columbia University Department of statistics.
- Gerster M, Waterstraat G, Litvak V, Lehnertz K, Schnitzler A, Florin E, Curio G, Nikulin V, 2021. Separating Neural Oscillations from Aperiodic 1/f activity: Challenges and Recommendations. *bioRxiv* (p. 2021.10.15.464483) doi: 10.1101/2021.10.15.464483.
- Gigerenzer G, 2008. Why heuristics work. *Perspect. Psychol. Sci.* 3 (1), 20–29. doi: 10.1111/j.1745-6916.2008.00058.x. [PubMed: 26158666]
- Gorgolewski KJ, Alfaro-Almagro F, Auer T, Bellec P, Capot M, Chakravarty MM, Churchill NW, Cohen AL, Craddock RC, Devenyi GA, Eklund A, Esteban O, Flandin G, Ghosh SS, Guntupalli JS, Jenkinson M, Keshavan A, Kiar G, Liem F, ... Poldrack RA, 2017. BIDS apps: improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. *PLoS Comput. Biol.* 13 (3), e1005209. doi: 10.1371/journal.pcbi.1005209. [PubMed: 28278228]
- Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, Flandin G, Ghosh SS, Glatard T, Halchenko YO, Handwerker DA, Hanke M, Keator D, Li X, Michael Z, Maumet C, Nichols BN, Nichols TE, Pellman J, ... Poldrack RA, 2016. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci. Data* 3, 160044. doi: 10.1038/sdata.2016.44. [PubMed: 27326542]
- Gorgolewski KJ, Poldrack RA, 2016. A practical guide for improving transparency and reproducibility in neuroimaging research. *PLoS Biol.* 14 (7), e1002506. doi: 10.1371/journal.pbio.1002506. [PubMed: 27389358]
- Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Goj R, Jas M, Brooks T, Parkkonen L, Hämäläinen M, 2013. MEG and EEG data analysis with MNE-python. *Front. Neurosci.* 7, 1–13. doi: 10.3389/fnins.2013.00267, 7 DEC. [PubMed: 23386807]
- Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Parkkonen L, Hämäläinen M, 2014. MNE software for processing MEG and EEG data. *NeuroImage* 86, 446–460. doi: 10.1016/j.neuroimage.2013.10.027. [PubMed: 24161808]
- Groppe DM, Urbach TP, Kutas M, 2011. Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. *Psychophysiology* 48 (12), 1711–1725. doi: 10.1111/j.1469-8986.2011.01273.x. [PubMed: 21895683]
- Gross J, Baillet S, Barnes GR, Henson RN, Hillebrand A, Jensen O, Jerbi K, Litvak V, Maess B, Oostenveld R, Parkkonen L, Taylor JR, van Wassenhove V, Wibrall M, Schoffelen JM, 2013. Good practice for conducting and reporting MEG research. *Neuroimage* 65, 349–363. doi: 10.1016/j.neuroimage.2012.10.001. [PubMed: 23046981]
- Hämäläinen MS, Ilmoniemi RJ, 1994. Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32 (1), 35–42. doi: 10.1007/BF02512476. [PubMed: 8182960]
- Handy TC, 2005. *Event-Related Potentials: A Methods Handbook*. MIT Press.
- Hari R, Baillet S, Barnes G, Burgess R, Forss N, Gross J, Hämäläinen M, Jensen O, Kakigi R, Mauguière F, Nakasato N, Puce A, Romani G-L., Schnitzler A., Taulu S., 2018. IFCN-endorsed practical guidelines for clinical magnetoencephalography (MEG). *Clin. Neurophysiol.* 129 (8), 1720–1747. doi: 10.1016/j.clinph.2018.03.042. [PubMed: 29724661]
- Hari R, Puce A, 2017. *MEG-EEG Primer*, 1st ed. Oxford University Press doi: 10.1093/med/9780190497774.001.0001.
- Hartmann KG, Schirrmeister RT, Ball T, 2018. EEG-GAN: Generative adversarial Networks For Electroencephalographic (EEG) Brain Signals. *ArXiv* 1806.01875 [Cs, Eess, q-Bio, Stat] doi: 10.48550/arXiv.1806.01875.
- Haselager P, 2013. Did i do that? Brain-computer interfacing and the sense of agency. *Minds Mach.* 23 (3), 405–418. doi: 10.1007/s11023-012-9298-7.
- Hassabis D, Kumaran D, Summerfield C, Botvinick M, 2017. Neuroscience-inspired artificial intelligence. *Neuron* 95 (2), 245–258. doi: 10.1016/j.neuron.2017.06.011. [PubMed: 28728020]

- Haufe S, Meinecke F, Görgen K, Dähne S, Haynes J-D., Blankertz B., Bießmann F., 2014. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage* 87, 96–110. doi: 10.1016/j.neuroimage.2013.10.067. [PubMed: 24239590]
- Hayasaka S, Nichols TE, 2003. Validating cluster size inference: random field and permutation methods. *Neuroimage* 20 (4), 2343–2356. doi: 10.1016/j.neuroimage.2003.08.003. [PubMed: 14683734]
- Hedderich DM, Eickhoff SB, 2021. Machine learning for psychiatry: getting doctors at the black box? *Mol. Psychiatry* 26 (1), 23–25. doi: 10.1038/s41380-020-00931-z. [PubMed: 33173196]
- Henrich J, Heine SJ, Norenzayan A, 2010. Most people are not WEIRD. *Nature* 466 (7302), 29–29. [PubMed: 20595995]
- Henson RN, Abdulrahman H, Flandin G, Litvak V, 2019. Multimodal Integration of M/EEG and f/MRI Data in SPM12. *Front. Neurosci.* 13. <https://www.frontiersin.org/article/10.3389/fnins.2019.00300>.
- Hinss MF, Somon B, Dehais F, Roy RN, 2021. Open EEG datasets for passive brain-computer interface applications: lacks and perspectives. In: *Proceedings of the 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 686–689. doi: 10.1109/NER49283.2021.9441214.
- Hintze JL, Nelson RD, 1998. Violin plots: a box plot-density trace synergism. *Am. Stat.* 52 (2), 181–184. doi: 10.1080/00031305.1998.10480559.
- Hoffmann S, Schönbrodt F, Elsas R, Wilson R, Strasser U, Boulesteix AL, 2021. The multiplicity of analysis strategies jeopardizes replicability: lessons learned across disciplines. *Royal Society Open Science* 8 (4), 201925. doi: 10.1098/rsos.201925. [PubMed: 33996122]
- Holdgraf C, Appelhoff S, Bickel S, Bouchard K, D'Ambrosio S, David O, Devinsky O, Dichter B, Flinker A, Foster BL, Gorgolewski KJ, Groen I, Groppe D, Gunduz A, Hamilton L, Honey CJ, Jas M, Knight R, Lachaux JP, ... Hermes D, 2019. IIEG-BIDS, extending the Brain Imaging Data Structure specification to human intracranial electrophysiology. *Sci. Data* 6 (1), 102. doi: 10.1038/s41597-019-0105-7. [PubMed: 31239438]
- Hume D, 1739. *A Treatise of Human Nature. A Treatise of Human Nature..* Printed for John Noon. Printed for John Noon.
- Hyman SE, 2011. Cognitive enhancement: promises and perils. *Neuron* 69 (4), 595–598. doi: 10.1016/j.neuron.2011.02.012. [PubMed: 21338872]
- Ienca M, Andorno R, 2017. Towards new human rights in the age of neuroscience and neurotechnology. *Life Sci. Soc. Policy* 13 (1), 5. doi: 10.1186/s40504-017-0050-1. [PubMed: 28444626]
- Ienca M, Haselager P, 2016. Hacking the brain: brain-computer interfacing technology and the ethics of neurosecurity. *Ethics Inf. Technol.* 18 (2), 117–129. doi: 10.1007/s10676-016-9398-9.
- Ienca M, Haselager P, Emanuel EJ, 2018. Brain leaks and consumer neurotechnology. *Nat. Biotechnol.* 36 (9), 805–810. doi: 10.1038/nbt.4240. [PubMed: 30188521]
- Ilmoniemi RJ, Sarvas J, 2019. *Brain Signals: Physics and Mathematics of MEG and EEG.* Mit Press.
- Ioannidis JPA, 2005. Why most published research findings are false. *PLoS Med.* 2 (8), e124. doi: 10.1371/journal.pmed.0020124. [PubMed: 16060722]
- Jaiswal A, Nenonen J, Stenroos M, Gramfort A, Dalal SS, Westner BU, Litvak V, Mosher JC, Schoffelen JM, Witton C, Oostenveld R, Parkkonen L, 2020. Comparison of beamformer implementations for MEG source localization. *Neuroimage* 216, 116797. doi: 10.1016/j.neuroimage.2020.116797. [PubMed: 32278091]
- Jas M, Larson E, Engemann DA, Leppäkangas J, Taulu S, Hämäläinen M, Gramfort A, 2018. A reproducible MEG/EEG group study with the MNE software: recommendations, quality assessments, and good practices. *Front. Neurosci.* 12, 530. doi: 10.3389/fnins.2018.00530. [PubMed: 30127712]
- Kahneman D, 2013. *Thinking, Fast and Slow*, 1st ed. Farrar, Straus and Giroux.
- Kakisaka Y, Kubota Y, Wang ZI, Piao Z, Mosher JC, Gonzalez-Martinez J, Jin K, Alexopoulos AV, Burgess RC, 2012. Use of simultaneous depth and MEG recording may provide complementary information regarding the epileptogenic region. *Epileptic Disord.* 14 (3), 298–303. doi: 10.1684/epd.2012.0517. [PubMed: 22940092]

- Kane N, Acharya J, Beniczky S, Caboclo L, Finnigan S, Kaplan PW, Shibasaki H, Pressler R, van Putten MJAM, 2017. A revised glossary of terms most commonly used by clinical electroencephalographers and updated proposal for the report format of the EEG findings. Revision 2017. *Clin. Neurophysiol. Pract.* 2, 170–185. doi: 10.1016/j.cnp.2017.07.002. [PubMed: 30214992]
- Kappenman ES, Farrens JL, Zhang W, Stewart AX, Luck SJ, 2021. ERP CORE: an open resource for human event-related potential research. *Neuroimage* 225, 117465. doi: 10.1016/j.neuroimage.2020.117465. [PubMed: 33099010]
- Kappenman ES, Luck SJ, 2010. The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology* doi: 10.1111/j.1469-8986.2010.01009.x.
- Kappenman ES, Luck SJ, 2016. Best practices for event-related potential research in clinical populations. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 1 (2), 110–115. doi: 10.1016/j.bpsc.2015.11.007. [PubMed: 27004261]
- Keil A, Debener S, Gratton G, Junghöfer M, Kappenman ES, Luck SJ, Luu P, Miller GA, Yee CM, 2014. Committee report: publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography: guidelines for EEG and MEG. *Psychophysiology* 51 (1), 1–21. doi: 10.1111/psyp.12147. [PubMed: 24147581]
- Kellmeyer P, 2018. Big brain data: on the responsible use of brain data from clinical and consumer-directed neurotechnological devices. *Neuroethics* doi: 10.1007/s12152-018-9371-x.
- Kiebel SJ, Garrido MI, Moran RJ, Friston KJ, 2008. Dynamic causal modelling for EEG and MEG. *Cogn. Neurodyn.* 2 (2), 121–136. doi: 10.1007/s11571-008-9038-0. [PubMed: 19003479]
- Kincaid H, Dupré J, Wylie A, 2007. *Value-Free Science?: Ideals and Illusions*. Oxford University Press doi: 10.1093/acprof:oso/9780195308969.001.0001.
- King JR, Dehaene S, 2014. Characterizing the dynamics of mental representations: the temporal generalization method. *Trends Cogn. Sci.* doi: 10.1016/j.tics.2014.01.002.
- Kirchberger K, Hummel C, Stefan H, 1998. Postoperative multichannel magnetoencephalography in patients with recurrent seizures after epilepsy surgery. *Acta Neurol. Scand.* 98 (1), 1–7. doi: 10.1111/j.1600-0404.1998.tb07370.x. [PubMed: 9696520]
- Klein G, 2007. Performing a project premortem. *Harv. Bus. Rev.* 85 (9), 18–19.
- Koessler L, Cecchin T, Colnat-Coulbois S, Vignal J-P, Jonas J, Vespignani H., Ramantani G., Maillard LG., 2015. Catching the invisible: mesial temporal source contribution to simultaneous EEG and SEEG recordings. *Brain Topogr.* 28 (1), 5–20. doi: 10.1007/s10548-014-0417-z. [PubMed: 25432598]
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI, 2009. Circular analysis in systems neuroscience – the dangers of double dipping. *Nat. Neurosci.* 12 (5), 535–540. doi: 10.1038/nn.2303. [PubMed: 19396166]
- Krol LR, Haselager P, Zander TO, 2020. Cognitive and affective probing: a tutorial and review of active learning for neuroadaptive technology. *J. Neural Eng.* 17 (1), 012001. doi: 10.1088/1741-2552/ab5bb5. [PubMed: 31770724]
- Krol LR, Pawlitzki J, Lotte F, Gramann K, Zander TO, 2018. SEREEGA: simulating event-related EEG activity. *J. Neurosci. Methods* 309, 13–24. doi: 10.1016/j.jneumeth.2018.08.001. [PubMed: 30114381]
- Lakens D, Caldwell AR, 2021. Simulation-based power analysis for factorial analysis of variance designs. *Adv. Methods Pract. Psychol. Sci.* 4 (1), 2515245920951503. doi: 10.1177/2515245920951503.
- Lancaster AK, Thessen AE, Virapongse A, 2018. A new paradigm for the scientific enterprise: nurturing the ecosystem. *F1000Research* 7, 803. doi: 10.12688/f1000research.15078.1. [PubMed: 30345023]
- Larson-Prior LJ, Oostenveld R, Della Penna S, Michalareas G, Prior F, Babajani-Feremi A, Schoffelen JM, Marzetti L, de Pasquale F, Di Pompeo F, Stout J, Woolrich M, Luo Q, Bucholz R, Fries P, Pizzella V, Romani GL, Corbetta M, Snyder AZ, 2013. Adding dynamics to the human connectome project with MEG. *Neuroimage* 80, 190–201. doi: 10.1016/j.neuroimage.2013.05.056. [PubMed: 23702419]

- Lehongre K, Lambrecq V, Whitmarsh S, Frazzini V, Cousyn L, Soleil D, ... & Navarro V (2022). Long-term deep intracerebral microelectrode recordings in patients with drug-resistant epilepsy: Proposed guidelines based on 10-year experience. *NeuroImage*, 254, 119116. doi: 10.1016/j.neuroimage.2022.119116. [PubMed: 35318150]
- Lemm S, Blankertz B, Dickhaus T, Müller K-R., 2011. Introduction to machine learning for brain imaging. *Neuroimage* 56 (2), 387–399. doi: 10.1016/j.neuroimage.2010.11.004. [PubMed: 21172442]
- Levitis E, van Praag CDG, Gau R, Heunis S, DuPre E, Kiar G, Bottenhorn KL, Glatard T, Nikolaidis A, Whitaker KJ, Mancini M, Niso G, Afyouni S, Alonso-Ortiz E, Appelhoff S, Arnatkeviciute A, Atay SM, Auer T, Baracchini G, ... Maumet C, 2021. Centering inclusivity in the design of online conferences —An OHBM–open science perspective. *GigaScience* 10 (8). doi: 10.1093/gigascience/giab051.
- Libet B, Gleason CA, Wright EW, Pearl DK, 1983. Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain* 106 (Pt 3), 623–642. doi: 10.1093/brain/106.3.623, *A Journal of Neurology*. [PubMed: 6640273]
- Litvak V, Mattout J, Kiebel S, Phillips C, Henson R, Kilner J, Barnes G, Oostenveld R, Daunizeau J, Flandin G, Penny W, Friston K, 2011. EEG and MEG data analysis in SPM8. *Comput. Intell. Neurosci.* 2011, 852961. doi: 10.1155/2011/852961. [PubMed: 21437221]
- Llorens A, Tzovara A, Bellier L, Bhaya-Grossman I, Bidet-Caulet A, Chang WK, Cross ZR, Dominguez-Faus R, Flinker A, Fonken Y, Gorenstein MA, Holdgraf C, Hoy CW, Ivanova MV, Jimenez RT, Jun S, Kam JWY, Kidd C, Marcelle E, ... Dronkers NF, 2021. Gender bias in academia: a lifetime problem that needs solutions. *Neuron* 109 (13), 2047–2074. doi: 10.1016/j.neuron.2021.06.002. [PubMed: 34237278]
- Lopez-Calderon J, Luck SJ, 2014. ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front. Hum. Neurosci.* 8, 213. doi: 10.3389/fnhum.2014.00213. [PubMed: 24782741]
- Lotte F, Bougrain L, Cichocki A, Clerc M, Congedo M, Rakotomamonjy A, Yger F, 2018. A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *J. Neural Eng.* 15 (3), 031005. doi: 10.1088/1741-2552/aab2f2. [PubMed: 29488902]
- Luck SJ, 2005. *An Introduction to the Event-Related Potential Technique*. MIT Press.
- Luck SJ, 2014. *An Introduction to the Event-Related Potential Technique*, 2nd ed. The MIT Press.
- Luck SJ, Gaspelin N, 2017. How to get statistically significant effects in any ERP experiment (and why you should not). *Psychophysiology* 54 (1), 146–157. doi: 10.1111/psyp.12639. [PubMed: 28000253]
- Luck SJ, Stewart AX, Simmons AM, Rhemtulla M, 2021. Standardized measurement error: a universal metric of data quality for averaged event-related potentials. *Psychophysiology* e13793. doi: 10.1111/psyp.13793. [PubMed: 33782996]
- Maestre FT, 2019. Ten simple rules towards healthier research labs. *PLoS Comput. Biol.* 15 (4), e1006914. doi: 10.1371/journal.pcbi.1006914. [PubMed: 30973866]
- Makeig S, Jung TP, Bell AJ, Ghahremani D, Sejnowski TJ, 1997. Blind separation of auditory event-related brain responses into independent components. *Proc. Natl. Acad. Sci.* 94 (20), 10979–10984. doi: 10.1073/pnas.94.20.10979. [PubMed: 9380745]
- Maris E, Oostenveld R, 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164 (1), 177–190. doi: 10.1016/j.jneumeth.2007.03.024. [PubMed: 17517438]
- Markiewicz CJ, Gorgolewski KJ, Feingold F, Blair R, Halchenko YO, Miller E, Hardcastle N, Wexler J, Esteban O, Goncavles M, Jwa A, Poldrack R, 2021. The OpenNeuro resource for sharing of neuroscience data. *Elife*. doi: 10.7554/eLife.71774.
- Markowitz F, 2015. Five selfish reasons to work reproducibly. *Genome Biol.* 16 (1), 274. doi: 10.1186/s13059-015-0850-7. [PubMed: 26646147]
- Martínez-Cancino R, Delorme A, Truong D, Artoni F, Kreutz-Delgado K, Sivagnanam S, Yoshimoto K, Majumdar A, Makeig S, 2021. The open EEGLAB portal Interface: high-performance computing with EEGLAB. *Neuroimage* 224, 116778. doi: 10.1016/j.neuroimage.2020.116778. [PubMed: 32289453]

- Martinovic I, Davies D, Frank M, Perito D, Ros T, Song D, 2012. On the feasibility of side-channel attacks with brain-computer interfaces. In: Proceedings of the 21st USENIX Security Symposium (USENIX Security 12), pp. 143–158.
- Masicampo EJ, Lalonde DR, 2012. A peculiar prevalence of p values just below .05. *Q. J. Exp. Psychol.* 65 (11), 2271–2279. doi: 10.1080/17470218.2012.711335.
- McKiernan EC, Bourne PE, Brown CT, Buck S, Kenall A, Lin J, McDougall D, Nosek BA, Ram K, Soderberg CK, Spies JR, Thaney K, Updegrove A, Woo KH, Yarkoni T, 2016. How open science helps researchers succeed. *ELife* 5, e16800. doi: 10.7554/eLife.16800. [PubMed: 27387362]
- Mecacci G, Haselager P, 2019. Identifying criteria for the evaluation of the implications of brain reading for mental privacy. *Sci. Eng. Ethics* 25 (2), 443–461. doi: 10.1007/s11948-017-0003-3. [PubMed: 29247306]
- Melnik A, Legkov P, Izdebski K, Kärcher SM, Hairston WD, Ferris DP, König P, 2017. Systems, subjects, sessions: to what extent do these factors influence EEG data? *Front. Hum. Neurosci.* 11. doi: 10.3389/fnhum.2017.00150.
- Meyer M, Lamers D, Kayhan E, Hunnius S, Oostenveld R, 2021. Enhancing reproducibility in developmental EEG research: BIDS, cluster-based permutation tests, and effect sizes. *Dev. Cogn. Neurosci.* 52, 101036. doi: 10.1016/j.dcn.2021.101036. [PubMed: 34801856]
- Miller G, 2006. A scientist's nightmare: software problem leads to five retractions. *Science* 314 (5807), 1856–1857. doi: 10.1126/science.314.5807.1856, (New York, N.Y.). [PubMed: 17185570]
- Moses DA, Leonard MK, Makin JG, Chang EF, 2019. Real-time decoding of question-and-answer speech dialogue using human cortical activity. *Nat. Commun.* 10 (1), 3096. doi: 10.1038/s41467-019-10994-4. [PubMed: 31363096]
- Mosher JC, Funke ME, 2020. Towards best practices in clinical magnetoencephalography: patient preparation and data acquisition. *J. Clin. Neurophysiol.* 37 (6), 498–507. doi: 10.1097/WNP.0000000000000542, Official Publication of the American Electroencephalographic Society. [PubMed: 33165223]
- Nave K, Hannon E, & Snyder JS (2020, September 23). Replication and extension of nozaradan, peretz, missal and mouraux (2011). osf.io/rpvde
- Musk E, 2019. An integrated brain-machine interface platform with thousands of channels. *Journal of medical Internet research* 21 (10), e16194. doi: 10.2196/16194. [PubMed: 31642810]
- Neymotin SA, Daniels DS, Caldwell B, McDougal RA, Carnevale NT, Jas M, Moore CI, Hines ML, Hämäläinen M, Jones SR, 2019. Human Neocortical Neurosolver (HNN): a new software tool for interpreting the cellular and network origin of human MEG/EEG data. *BioRxiv* 740597. doi: 10.1101/740597.
- Nichols TE, Das S, Eickhoff SB, Evans AC, Glatard T, Hanke M, Kriegeskorte N, Milham MP, Poldrack RA, Poline JB, Proal E, Thirion B, Van Essen DC, White T, Yeo BTT, 2017. Best practices in data analysis and sharing in neuroimaging using MRI. *Nat. Neurosci.* 20 (3), 299–303. doi: 10.1038/nn.4500. [PubMed: 28230846]
- Nichols T, Hayasaka S, 2003. Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat. Methods Med. Res.* 12 (5), 419–446. doi: 10.1191/0962280203sm341ra. [PubMed: 14599004]
- Nieuwland MS, Politzer-Ahles S, Heyselaar E, Segaert K, Darley E, Kazanina N, Von Grebmer Zu Wolfsturn., Bartolozzi F, Kogan V, Ito A., Mézière D., Barr DJ., Rousselet GA., Ferguson HJ., Busch-Moreno S., Fu., Tuomainen J., Kulakova E., Husband EM., ... Huettig., 2018. Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife* 7, e33468. doi: 10.7554/eLife.33468. [PubMed: 29631695]
- Niso G, Gorgolewski KJ, Bock E, Brooks TL, Flandin G, Gramfort A, Henson RN, Jas M, Litvak V, Moreau J T., Oostenveld R., Schoffelen J-M., Tadel F., Wexler J., Baillet S., 2018. MEG-BIDS, the brain imaging data structure extended to magnetoencephalography. *Sci. Data* 5 (1), 180110. doi: 10.1038/sdata.2018.110. [PubMed: 29917016]
- Niso G, Rogers C, Moreau JT, Chen LY, Madjar C, Das S, Bock E, Tadel F, Evans AC, Jolicoeur P, Baillet S, 2016. OMEGA: the open MEG archive. *Neuroimage* 124, 1182–1187. doi: 10.1016/j.neuroimage.2015.04.028. [PubMed: 25896932]

- Niso G, Tadel F, Bock E, Cousineau M, Santos A, Baillet S, 2019. Brainstorm pipeline analysis of resting-state data from the open MEG archive. *Front. Neurosci.* 13, 284. doi: 10.3389/fnins.2019.00284. [PubMed: 31024228]
- Niso G, Botvinik-Nezer R, Appelhoff S, De La Vega A, Esteban O, Etzel JA, Rieger J, 2022. Open and reproducible neuroimaging: from study inception to publication. *OSF Preprints* doi: 10.31219/osf.io/pu5vb.
- Norgaard M, Matheson GJ, Hansen HD, Thomas A, Searle G, Rizzo G, Veronese M, Giacomel A, Yaqub M, Tonietto M, Funck T, Gillman A, Boniface H, Routier A, Dalenberg JR, Betthausen T, Feingold F, Markiewicz CJ, Gorgolewski KJ, ... Ganz M, 2021. PET-BIDS, an extension to the brain imaging data structure for positron emission tomography. *BioRxiv* doi: 10.1101/2021.06.16.448390, 2021.06.16.448390.
- Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, Buck S, Chambers CD, Chin G, Christensen G, Contestabile M, Dafoe A, Eich E, Freese J, Glennerster R, Goroff D, Green DP, Hesse B, Humphreys M, ... Yarkoni T, 2015. Promoting an open research culture. *Science* 348 (6242), 1422–1425. doi: 10.1126/science.aab2374. [PubMed: 26113702]
- Nosek BA, Ebersole CR, DeHaven AC, Mellor DT, 2018. The preregistration revolution. *Proc. Natl. Acad. Sci.* 115 (11), 2600–2606. doi: 10.1073/pnas.1708274114. [PubMed: 29531091]
- Nosek BA, Hardwicke TE, Moshontz H, Allard A, Corker KS, Dreber A, Fidler F, Hilgard J, Kline Struhl M, Nuijten MB, Rohrer JM, Romero F, Scheel AM, Scherer LD, Schönbrodt FD, Vazire S, 2022. Replicability, robustness, and reproducibility in psychological science. *Annu. Rev. Psychol.* 73 (1), 719–748. doi: 10.1146/annurev-psych-020821-114157. [PubMed: 34665669]
- Nosek BA, Spies JR, Motyl M, 2012. Scientific Utopia: II. restructuring incentives and practices to promote truth over publishability. *Perspect. Psychol. Sci.* 7 (6), 615–631. doi: 10.1177/1745691612459058. [PubMed: 26168121]
- Obeid I, Picone J, 2016. The temple university hospital EEG data corpus. *Front. Neurosci.* 10, 196. doi: 10.3389/fnins.2016.00196. [PubMed: 27242402]
- Olivet DM, Hajcak G, 2009. Reliability of error-related brain activity. *Brain Res.* 1284, 89–99. doi: 10.1016/j.brainres.2009.05.079. [PubMed: 19501071]
- Oostenveld R, Fries P, Maris E, Schoffelen JM, 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011 (3), 156869. doi: 10.1155/2011/156869. [PubMed: 21253357]
- Open Science Collaboration, 2015. Estimating the reproducibility of psychological science. *Science* 349 (6251). doi: 10.1126/science.aac4716, aac4716.
- Palva JM, Wang SH, Palva S, Zhigalov A, Monto S, Brookes MJ, Schoffelen JM, Jerbi K, 2018. Ghost interactions in MEG/EEG source space: a note of caution on inter-areal coupling measures. *Neuroimage* 173, 632–643. doi: 10.1016/j.neuroimage.2018.02.032. [PubMed: 29477441]
- Pardo MS, Patterson D, 2013. *Minds, Brains, and Law: The Conceptual Foundations of Law and Neuroscience*. Oxford University Press doi: 10.1093/acprof:oso/9780199812134.001.0001.
- Parvizi J, Kastner S, 2018. Promises and limitations of human intracranial electroencephalography. *Nat. Neurosci.* 21 (4), 474–483. doi: 10.1038/s41593-018-0108-2. [PubMed: 29507407]
- Paul M, Govaart GH, Schettino A, 2021. Making ERP research more transparent: guidelines for preregistration. *Int. J. Psychophysiol.* 164, 52–63. doi: 10.1016/j.ijpsycho.2021.02.016. [PubMed: 33676957]
- Pavlov YG, Adamian N, Appelhoff S, Arvaneh M, Benwell CSY, Beste C, Bland AR, Bradford DE, Bublatzky F, Busch NA, Clayson PE, Cruse D, Czeszumski A, Dreber A, Dumas G, Ehinger B, Giorgio G, He X, Hinojosa JA, ... Mushtaq F, 2021. #EEGManyLabs: investigating the replicability of influential EEG experiments. *Cortex* doi: 10.1016/j.cortex.2021.03.013.
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, 2011. Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Pernet C, Appelhoff S, Gorgolewski KJ, Flandin G, Phillips C, Delorme A, Oostenveld R, 2019. EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Sci. Data* 6 (1), 103. doi: 10.1038/s41597-019-0104-8. [PubMed: 31239435]
- OECD. Recommendation of the council on responsible innovation in neurotechnology.

- Pernet C, Garrido M, Gramfort A, Maurits N, Michel C, Pang E, Salmelin R, Schoffelen JM, Valdes-Sosa PA, Puce A, 2018. Best practices in data analysis and sharing in neuroimaging using MEEG. OSF Preprints doi: 10.31219/osf.io/a8dhx.
- Pernet C, Garrido MI, Gramfort A, Maurits N, Michel CM, Pang E, Salmelin R, Schoffelen JM, Valdes-Sosa PA, Puce A, 2020. Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research. *Nat. Neurosci.* 23 (12), 1473–1483. doi: 10.1038/s41593-020-00709-0. [PubMed: 32958924]
- Pernet C, Martinez-Cancino R, Truong D, Makeig S, Delorme A, 2021. From BIDS-formatted EEG data to sensor-space group results: a fully reproducible workflow with EEGLAB and LIMO EEG. *Front. Neurosci.* 14, 610388. doi: 10.3389/fnins.2020.610388. [PubMed: 33519362]
- Picton TW, Bentin S, Berg P, Donchin E, Hillyard SA, Johnson R, Miller GA, Ritter W, Ruchkin DS, Rugg MD, Taylor MJ, 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37 (2), 127–152. doi: 10.1111/1469-8986.3720127. [PubMed: 10731765]
- Pivik RT, Broughton RJ, Coppola R, Davidson RJ, Fox N, Nuwer MR, 1993. Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts. *Psychophysiology* 30 (6), 547–558. doi: 10.1111/j.1469-8986.1993.tb02081.x. [PubMed: 8248447]
- Piwowar HA, Vision TJ, 2013. Data reuse and the open data citation advantage. *PeerJ* 1, e175. doi: 10.7717/peerj.175. [PubMed: 24109559]
- Pizzo F, Roehri N, Medina Villalon S, Trébuchon A, Chen S, Lagarde S, Carron R, Gavaret M, Giusiano B, McGonigal A, Bartolomei F, Badier JM, Béнар CG, 2019. Deep brain activities can be detected with magnetoencephalography. *Nat. Commun.* 10 (1), 971. doi: 10.1038/s41467-019-08665-5. [PubMed: 30814498]
- Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munafò MR, Nichols TE, Poline J-B., Vul E., Yarkoni T., 2017. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat. Rev. Neurosci.* 18 (2), 115–126. doi: 10.1038/nrn.2016.167. [PubMed: 28053326]
- Poldrack RA, Whitaker K, Kennedy D, 2020. Introduction to the special issue on reproducibility in neuroimaging. *Neuroimage* 218, 116357. doi: 10.1016/j.neuroimage.2019.116357. [PubMed: 31733374]
- Pontifex MB, Scudder MR, Brown ML, O’Leary KC, Wu CT, Themanson JR, Hillman CH, 2010. On the number of trials necessary for stabilization of error-related brain activity across the life span. *Psychophysiology* 47 (4), 767–773. doi: 10.1111/j.1469-8986.2010.00974.x. [PubMed: 20230502]
- Popov T, Oostenveld R, Schoffelen JM, 2018. FieldTrip made easy: an analysis protocol for group analysis of the auditory steady state brain response in time, frequency, and space. *Front. Neurosci.* 12. <https://www.frontiersin.org/article/10.3389/fnins.2018.00711>.
- Proctor RN, 1991. *Value-Free Science? Purity and Power in Modern Knowledge*. Harvard University Press.
- Puolivali T, Palva S, Palva JM, 2020. Influence of multiple hypothesis testing on reproducibility in neuroimaging research: a simulation study and Python-based software. *J. Neurosci. Methods* 337, 108654. doi: 10.1016/j.jneumeth.2020.108654. [PubMed: 32114144]
- Quinn AJ, Lopes-Dos-Santos V, Dupret D, Nobre AC, & Woolrich MW (2021). EMD: Empirical Mode Decomposition and Hilbert-Huang Spectral Analyses in Python. *Journal of open source software*, 6(59), 2977. 10.21105/joss.02977. [PubMed: 33855259]
- Rae CL, Farley M, Jeffery KJ, Urai AE, 2022. Climate crisis and ecological emergency: Why they concern (neuro) scientists, and what we can do. *Brain and Neuroscience Advances* 6. doi: 10.1177/23982128221075430.
- Renkewitz F, Heene M, 2019. The replication crisis and open science in psychology. *Zeitschrift Für Psychol.* 227 (4), 233–236. doi: 10.1027/2151-2604/a000389.
- Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, Aviles-Rivero AI, Etmann C, McCague C, Beer L, Weir-McCall JR, Teng Z, Gkrania-Klotsas E, Rudd JHF, Sala E, Schönlieb CB, 2021. Common pitfalls and recommendations for using machine learning to detect and

- prognosticate for COVID-19 using chest radiographs and CT scans. *Nat. Mach. Intell.* 3 (3), 199–217. doi: 10.1038/s42256-021-00307-0.
- Rodu J, Kafadar K, 2021. The q–q Boxplot. *J. Comput. Graph. Stat.* 0 (0), 1–14. doi: 10.1080/10618600.2021.1938586.
- Rosenfeld JP, Labkovsky E, Winograd M, Lui MA, Vandenboom C, Chedid E, 2008. The complex trial protocol (CTP): a new, countermeasure-resistant, accurate, P300-based method for detection of concealed information. *Psychophysiology* 45 (6), 906–919. doi: 10.1111/j.1469-8986.2008.00708.x. [PubMed: 18823418]
- Rosenthal R, 1979. The file drawer problem and tolerance for null results. *Psychol. Bull.* 86 (3), 638. doi: 10.1037/0033-2909.86.3.638.
- Roy Y, Banville H, Albuquerque I, Gramfort A, Falk TH, Faubert J, 2019. Deep learning-based electroencephalography analysis: a systematic review. *J. Neural Eng.* 16 (5), 051001. doi: 10.1088/1741-2552/ab260c. [PubMed: 31151119]
- Rubin M, 2020. Does preregistration improve the credibility of research findings? *Quant. Methods Psychol.* 16 (4), 376–390. doi: 10.20982/tqmp.16.4.p376.
- Rust J, Golombok S, 2014. *Modern Psychometrics: The Science of Psychological Assessment*, 3rd ed. Routledge doi: 10.4324/9781315787527.
- Salo P, Heikkinen H, 2018. Slow science: research and teaching for sustainable praxis. *Confero* 6 (1). doi: 10.3384/confero.2001-4562.181130.
- Sandre 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. doi: 10.1016/j.ijpsycho.2020.06.016. [PubMed: 32653551]
- Santiuste M, Nowak R, Russi A, Tarancon T, Oliver B, Ayats E, Scheler G, Graetz G, 2008. Simultaneous magnetoencephalography and intracranial EEG registration: technical and clinical aspects. *J. Clin. Neurophysiol.* 25 (6), 331–339. doi: 10.1097/WNP.0b013e31818e7913. [PubMed: 18997623]
- Sanz Leon P, Knock SA, Woodman MM, Domide L, Mersmann J, McIntosh AR, Jirsa V, 2013. The virtual brain: a simulator of primate brain network dynamics. *Front. Neuroinform.* 7 doi: 10.3389/fninf.2013.00010.
- Sassenhagen J, Draschkow D, 2019. Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. *Psychophysiology* 56 (6), e13335. doi: 10.1111/psyp.13335. [PubMed: 30657176]
- Scheel AM, Schijen MRMJ, Lakens D, 2021. An excess of positive results: comparing the standard psychology literature with registered reports. *Adv. Methods Pract. Psychol. Sci.* 4 (2), 25152459211007468. doi: 10.1177/25152459211007467.
- Schönbrodt FD, Wagenmakers EJ, 2018. Bayes factor design analysis: planning for compelling evidence. *Psychon. Bull. Rev.* 25 (1), 128–142. doi: 10.3758/s13423-017-1230-y. [PubMed: 28251595]
- Schreiwies C, Volle E, Durr A, Auffret A, Delarasse C, George N, Dumont M, Hassan BA, Renier N, Rosso C, Thiebaut de Schotten M, Burguière E, Zujovic V, 2019. A neuroscientific approach to increase gender equality. *Nat. Hum. Behav.* 3 (12), 1238–1239. doi: 10.1038/s41562-019-0755-7. [PubMed: 31570760]
- Schrouff J, Pischetta D, Genon S, Fryns G, Pinho AL, Vassena E, Liuzzi AG, Ferreira FS, 2019. Gender bias in (neuro)science: facts, consequences, and solutions. *Eur. J. Neurosci.* 50 (7), 3094–3100. doi: 10.1111/ejn.14397. [PubMed: 30811680]
- Schultze-Kraft M, Birman D, Rusconi M, Allefeld C, Gørgen K, Dähne S, Blankertz B, Haynes JD, 2016. The point of no return in vetoing self-initiated movements. *Proc. Natl. Acad. Sci.* 113 (4), 1080–1085. doi: 10.1073/pnas.1513569112. [PubMed: 26668390]
- Seeber M, Cantonas LM, Hoevens M, Sesia T, Visser-Vandewalle V, Michel CM, 2019. Subcortical electrophysiological activity is detectable with high-density EEG source imaging. *Nat. Commun.* 10 (1), 753. doi: 10.1038/s41467-019-08725-w. [PubMed: 30765707]
- Shafto M.a, Tyler LK., Dixon M., Taylor JR., Rowe JB., Cusack R., Calder AJ., Marslen-Wilson WD., Duncan J., Dalgleish T., Henson RN., Brayne C., Matthews FE., 2014. *The Cambridge*

centre for ageing and neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC Neurol.* 14, 204. doi: 10.1186/s12883-014-0204-1. [PubMed: 25412575]

- Shaw C, Ward L, 2014. Dark thoughts: Why Mental Illness is On the Rise in Academia March 6. *The Guardian*.
- Shrout PE, Rodgers JL, 2018. Psychology, science, and knowledge construction: broadening perspectives from the replication crisis. *Annu. Rev. Psychol.* 69 (1), 487–510. doi: 10.1146/annurev-psych-122216-011845. [PubMed: 29300688]
- Simmons AM, Luck SJ, 2020. Protocol for reducing COVID-19 transmission risk in EEG research. *Res. Sq.* doi: 10.21203/rs.3.pex-974/v2, rs.3.pex-974.
- Simmons JP, Nelson LD, Simonsohn U, 2011. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* 22 (11), 1359–1366. doi: 10.1177/0956797611417632. [PubMed: 22006061]
- Sinha SR, Sullivan L, Sabau D, San-Juan D, Dombrowski KE, Halford JJ, Hani AJ, Drislane FW, Stecker MM American Clinical Neurophysiology Society Guideline 1: Minimum Technical Requirements for Performing Clinical Electroencephalography. *J Clin Neurophysiol.* 2016 Aug;33(4):303–7. doi: 10.1097/WNP.0000000000000308. Erratum in: *J Clin Neurophysiol.* 2021 May 1;38(3):e16. [PubMed: 27482788]
- Sinnott-Armstrong W, 2011. *Lessons from Libet. Conscious Will and Responsibility In Conscious Will and Responsibility.* Oxford University Press.
- Sinnott-Armstrong W, Simmons C, 2021. Some common fallacies in arguments from M/EEG data. *Neuroimage* 245, 118725. doi: 10.1016/j.neuroimage.2021.118725. [PubMed: 34813968]
- Smaldino PE, McElreath R, 2016. The natural selection of bad science. *R. Soc. Open Sci.* 3 (9), 160384. doi: 10.1098/rsos.160384. [PubMed: 27703703]
- Smith SM, Nichols TE, 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44 (1), 83–98. doi: 10.1016/j.neuroimage.2008.03.061. [PubMed: 18501637]
- Soderberg CK, Errington TM, Schiavone SR, Bottesini J, Thorn FS, Vazire S, Esterling KM, Nosek BA, 2021. Initial evidence of research quality of registered reports compared with the standard publishing model. *Nat. Hum. Behav.* 5 (8), 990–997. doi: 10.1038/s41562-021-01142-4. [PubMed: 34168323]
- Sood G, Laohaprapanon S, 2018. Predicting race and ethnicity from the sequence of characters in a name. *ArXiv* doi: 10.48550/arXiv.1805.02109.
- Šoški A, Jovanovi V, Styles SJ, Kappenman ES, Kovic V, 2021. How to do better N400 studies: reproducibility, consistency and adherence to research standards in the existing literature. *Neuropsychol. Rev.* doi: 10.1007/s11065-021-09513-4.
- Šoški A, Kappenman ES, Styles SJ, Kovi V. Garden of forking paths in ERP research – effects of varying processing and analysis steps in an N400 experiment. *OSF*.
- Šoški A, Kovi V, Ke H, Styles SJ, 2020. ARTEM-IS: agreed Reporting Template for EEG Methodology -. International Standard. *OSF*. doi: 10.17605/OSF.IO/PVRN6.
- Steege S, Tuerlinckx F, Gelman A, Vanpaemel W, 2016. Increasing transparency through a multiverse analysis. *Perspect. Psychol. Sci.* 11 (5), 702–712. doi: 10.1177/1745691616658637. [PubMed: 27694465]
- Stengers I, 2013. *Another Science is Possible. A Manifesto For Slow Science.* Polity Press (Muecke S. Translation).
- Stenner T, Boulay C, Grivich M, Medine D, Kothe C, Grimm G, Xloem T, Biancarelli A, Mansencal B, Frey JK, Powell S, Clisson P, & phfix. 2021. *scn/liblsl: V1.15.1 - ASIO updated.* Zenodo. doi: 10.5281/zenodo.5415959.
- Strickland E, Gallucci M, 2022. First win for the neurorights campaign: chile plans to regulate all neurotech and ban the sale of brain data. *IEEE Spectr.* 59 (1), 26–58. doi: 10.1109/MSPEC.2022.9676352.
- Styles SJ, Kovic V, Ke H, Šoški A, 2020. ARTEM-IS Template for ERP - Draft for Consultation. *OSF*.

- Styles SJ, Kovic V, Ke H, Šoški A, 2021. Towards ARTEM-IS: an evidence-based agreed reporting template for EEG methodology - international standard. *Neuroimage* 245, 118721. doi: 10.1016/j.neuroimage.2021.118721. [PubMed: 34826594]
- Szucs D, Ioannidis JPA, 2017. Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. *PLoS Biol.* 15 (3), e2000797. doi: 10.1371/journal.pbio.2000797. [PubMed: 28253258]
- Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM, 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011, 879716. doi: 10.1155/2011/879716. [PubMed: 21584256]
- Tadel F, Bock E, Niso G, Mosher JC, Cousineau M, Pantazis D, Leahy RM, Baillet S, 2019. MEG/EEG group analysis with brainstorm. *Front. Neurosci.* 13, 76. doi: 10.3389/fnins.2019.00076. [PubMed: 30804744]
- Tallon-Baudry C, Bertrand O, 1999. Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci.* 3 (4), 151–162. doi: 10.1016/S1364-6613(99)01299-1. [PubMed: 10322469]
- Tanner D, Morgan-Short K, Luck SJ, 2015. How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition: high-pass filtering and artifactual ERP effects. *Psychophysiology* 52 (8), 997–1009. doi: 10.1111/psyp.12437. [PubMed: 25903295]
- Taylor JR, Williams N, Cusack R, Auer T, Shafto MA, Dixon M, Tyler LK, Cam-CAN X, Henson RN, 2015. The Cambridge centre for ageing and neuroscience (Cam-CAN) data repository: structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *Neuroimage* 144, 262–269. doi: 10.1016/j.neuroimage.2015.09.018. [PubMed: 26375206]
- Tibon R, Open Science Committee C, Henson R, 2018. Title TBA: revising the abstract submission process. *Trends Cogn. Sci.* 22 (4), 271–274. doi: 10.1016/j.tics.2018.01.008. [PubMed: 29482978]
- Tversky A, Kahneman D, 1974. Judgment under uncertainty: heuristics and biases. *Science* 185 (4157), 1124–1131. doi: 10.1126/science.185.4157.1124. [PubMed: 17835457]
- van Vliet M, 2019. Seven quick tips for data analysis pipelines in neuroimaging. *PLoS Comput. Biol.* 16 (3), e1007358. doi: 10.1371/journal.pcbi.1007358.
- Van Vliet M, Liljeström M, Aro S, Salmelin R, Kujala J, 2018. Analysis of functional connectivity and oscillatory power using DICS: from raw MEG data to group-level statistics in python. *Front. Neurosci.* 586. doi: 10.3389/fnins.2018.00586.
- Vazire S, 2017. Quality uncertainty erodes trust in science. *Collabra: Psychol.* 3 (1). doi: 10.1525/collabra.74.
- Vul E, Harris C, Winkielman P, Pashler H, 2009. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect. Psychol. Sci.* 4 (3), 274–290. doi: 10.1111/j.1745-6924.2009.01125.x. [PubMed: 26158964]
- Wang P, 2019. On defining artificial intelligence. *J. Artif. Gen. Intell.* 10 (2), 1–37. doi: 10.2478/jagi-2019-0002.
- Ware C, 2019. Information Visualization: Perception for Design. Morgan Kaufmann.
- Wen H, Liu Z, 2016. Separating fractal and oscillatory components in the power spectrum of neurophysiological signal. *Brain Topogr.* 29 (1), 13–26. doi: 10.1007/s10548-015-0448-0. [PubMed: 26318848]
- Wendling F, Benquet P, Bartolomei F, Jirsa V, 2016. Computational models of epileptiform activity. *J. Neurosci. Methods* 260, 233–251. doi: 10.1016/j.jneumeth.2015.03.027. [PubMed: 25843066]
- Westner BU, Dalal SS, Gramfort A, Litvak V, Mosher JC, Oostenveld R, Schoffelen JM, 2022. A unified view on beamformers for M/EEG source reconstruction. *Neuroimage* 246, 118789. doi: 10.1016/j.neuroimage.2021.118789. [PubMed: 34890794]
- Whiteford KL, Baltzell LS, Cooper JK, Irsik VC, Irvine A, Mesik J, Nolan T, Oakes B, Reed A, Schrlau AE, Van Hedger S, Bharadwaj H, Johnsrude I, Kidd G Jr., Lim SJ, Luebke AE, Maddox RK, Marvin EW, Perrachione T, ... Oxenham A 2020. Association of musical training with auditory and speech neural coding and perception. *OSF*. doi: 10.17605/OSF.IO/DUQ34.

- Widmann A, Schröger E, Maess B, 2015. Digital filter design for electrophysiological data – a practical approach. *J. Neurosci. Methods* 250, 34–46. doi: 10.1016/j.jneumeth.2014.08.002, Supplement C. [PubMed: 25128257]
- Wilkinson MD, Dumontier M, Aalbersberg., Appleton G., Axton M., Baak A., Blomberg N., Boiten J-W., da Silva Santos LB., Bourne PE., Bouwman J, Brookes AJ., Clark T., Crosas M., Dillo., Dumon O., Edmunds S., Evelo CT., Finkers R., ... Mons B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3 (1), 160018. doi: 10.1038/sdata.2016.18. [PubMed: 26978244]
- Wilson G, Aruliah DA, Brown CT, Hong NPC, Davis M, Guy RT, Haddock SHD, Huff KD, Mitchell IM, Plumbley MD, Waugh B, White EP, Wilson P, 2014. Best practices for scientific computing. *PLoS Biol.* 12 (1), e1001745. doi: 10.1371/journal.pbio.1001745. [PubMed: 24415924]
- Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM, 2002. Brain-computer interfaces for communication and control. *Clin. Neurophysiol* 113 (6). doi: 10.1016/S1388-2457(02)00057-3, 767–7. [PubMed: 12048038]
- Wolpaw J, Wolpaw EW, 2012. *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press.
- Xiong C, Van Weelden L, Franconeri S, 2020. The curse of knowledge in visual data communication. *IEEE Trans. Visual Comput. Graph.* 26 (10), 3051–3062. doi: 10.1109/TVCG.2019.2917689.
- Yael D, Vecht JJ, Bar-Gad I, 2018. Filter-based phase shifts distort neuronal timing information. *Eneuro* 5 (2). doi: 10.1523/ENEURO.0261-17.2018.
- Yarkoni T, Markiewicz C, de la Vega A, Gorgolewski K, Salo T, Halchenko Y, Mc-Namara Q, DeStasio K, Poline JB, Petrov D, Hayot-Sasson V, Nielson D, Carlin J, Kiar G, Whitaker K, DuPre E, Wagner A, Tirrell L, Jas M, ... Blair R, 2019. PyBIDS: python tools for BIDS datasets. *J. Open Source Softw.* 4 (40), 1294. doi: 10.21105/joss.01294. [PubMed: 32775955]
- Ylijoki OH, Mäntylä H, 2003. Conflicting time perspectives in academic work. *Time Soc.* 12 (1), 55–78. doi: 10.1177/0961463x03012001364.
- Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY, 2008. Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron* 58 (3), 429–441. doi: 10.1016/j.neuron.2008.03.027. [PubMed: 18466752]
- Zander TO, Kothe CA, 2011. Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *J. Neural Eng.* 8 (2), 025005. doi: 10.1088/1741-2560/8/2/025005. [PubMed: 21436512]
- Zander TO, Krol LR, Birbaumer NP, Gramann K, 2016. Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity. *Proc. Natl. Acad. Sci.* 113 (52), 14898–14903. doi: 10.1073/pnas.1605155114. [PubMed: 27956633]
- Zhang X, Yao L, Wang X, Monaghan J, McAlpine D, Zhang Y, 2021. A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers. *J. Neural Eng.* 18 (3), 031002. doi: 10.1088/1741-2552/abc902.
- Zhou D, Cornblath EJ, Stiso J, Teich EG, Dworkin JD, Blevins AS, Bassett DS, 2020. Gender Diversity Statement and Code Notebook v1.0. Zenodo doi: 10.5281/zenodo.3672110.
- Zurn P, Bassett DS, Rust NC, 2020. The Citation diversity statement: a practice of transparency, a way of life. *Trends Cogn. Sci.* 24 (9), 669–672. doi: 10.1016/j.tics.2020.06.009. [PubMed: 32762966]
- Stewart AX, Luck SJ, 2020. Standardized measurement error (SME) demo scripts. OSF <https://osf.io/a4huc/>.

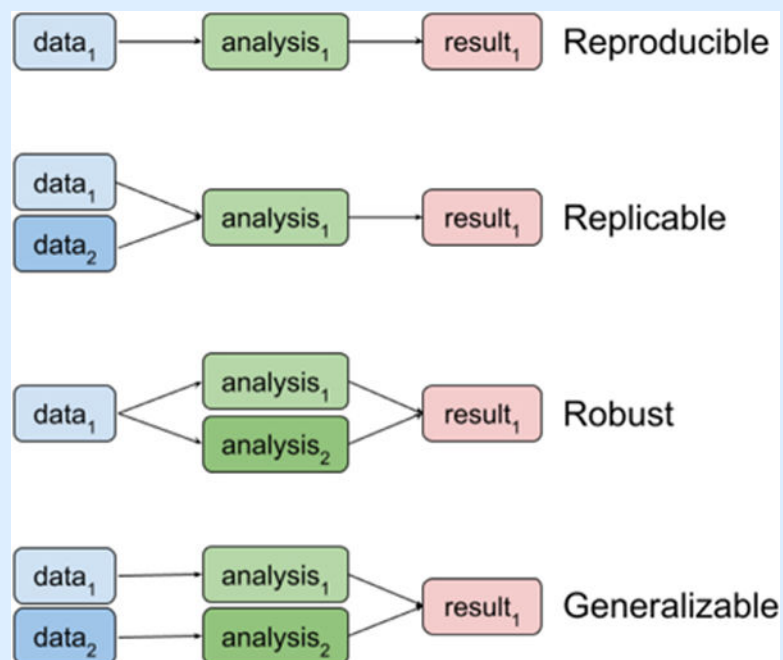
Box 1.**Evolving GSP**

GSP is under constant evolution. Current GSP may not be appropriate in the future, as new methods highlight limitations of current practice. Some examples are:

1. GSP for gamma-band oscillatory activity measurements has evolved dramatically after the discovery of a microsaccadic spike artifact in the EEG (Yuval-Greenberg et al., 2008) with energy in the gamma frequency range. This spike artifact led researchers to now routinely monitor saccadic eye movements prior to drawing conclusions regarding high-frequency brain activity recorded from the scalp.
2. Traditionally acceptable statistical practices in neuroimaging consisted of selecting regions of interest for further analysis with the same contrast that was used to report effect sizes in publications. This is now a well-known case of problematic circularity in data analysis (Kriegeskorte et al., 2009; Vul et al., 2009). Similarly, the failure to account for multiple comparisons in mass univariate data analysis (shown with irony by Bennett et al. (2009), or by Nichols and Hayasaka (2003)) has become completely unacceptable.
3. Wrong uses of powerful methods often emerge, especially if these are successful and widely used. For instance, the cluster-based correction for multiple comparisons (Maris and Oostenveld 2007) seems to imply a spatio-temporal extent of the effect under investigation and has been used for this purpose. However, this extent depends on an arbitrary threshold, and should not be used to localize effects in space or time, as has been reminded recently (Sassenhagen and Draschkow, 2019).
4. Similarly, increasing the number of factors in a statistical analysis (e.g., ANOVA) causes an exponential increase in the number of main effects and interactions being tested, which can dramatically increase the familywise Type I error rate (Cramer et al., 2016; Luck and Gaspelin, 2017). As a result, previous advice to include electrode/sensor as a factor in statistical analyses (Luck, 2005) has been replaced by advice to collapse across electrodes/sensors unless this factor is important for testing the scientific hypotheses (Luck, 2014).

Box 2**Reproducible research defined**

The recent reproducibility crisis in neuroimaging (Button et al., 2013; Poldrack et al., 2017) has highlighted the importance of conducting reproducible research (Poldrack et al., 2020) and stressed its value in GSP. While reproducibility is often used as an umbrella term in the literature, here we would like to provide more specific definitions for each of the dimensions of the process: reproducibility, replicability, robustness, and generalizability, as introduced by Community et al. (2019). Scientific work can be defined as *reproducible* when the same results can be systematically observed if the same analytic steps are carried out on the same dataset as the original study. Research is considered *replicable* if the same analysis performed on a different dataset (e.g., another group of participants) outputs qualitatively similar results. *Robust* scientific work occurs when the same dataset is used with a different analysis pipeline that nonetheless tackles the same research question as the original work (e.g., the original code in Python is translated to R, or when the same pipeline is run using different analysis toolboxes) and similar results are obtained. Finally, *generalizable* work combines new but similar datasets, with new but similar analysis pipelines, to produce similar results.



Box 3.**MEEG analysis software in brief**

Prior to computers becoming tools for MEEG analysis, “paper and pen” EEG data analysis was the norm. Acquisition hardware used to print time-courses on long rolls of graph paper or z-folding paper, from which amplitudes and even frequencies were measured with a ruler. Experimental data were noted in tables and statistical analyses performed by hand. Computers became popular for MEEG data analysis in the later parts of the 20th century. Numerous toolboxes were created to enable complex data analysis of high-density MEEG datasets at a speed that earlier users could not have dreamt of.

But with the computing power available today, it is easy to hunt for significant effects in the data, and ease-of-use can hide a method’s underlying complexity. Therefore, leading MEEG toolboxes have all developed strategies to safeguard against most common errors. Below, we examine some widely used toolboxes for MEEG analysis. All five toolboxes can run complete analysis pipelines, are free and open source, and methods originally developed with one toolbox are usually ported and usable in the others.

EEGLAB (Delorme and Makeig, 2004) is of the first, and currently the most-cited open-source toolboxes for EEG analysis. It was created at the Swartz Center for Computational Neuroscience at the University of California San Diego by Scott Makeig and Arnaud Delorme in the MATLAB scientific programming language (support for Octave starting in 2021). It is largely developed around independent component analysis (ICA), introduced to EEG by Makeig et al. (1997). ICA is a popular blind source separation method for both artifact rejection and data analysis (see also Martínez-Cancino et al., 2021). EEGLAB can incorporate independent components as virtual channels for both standard and ICA-specific source-level analyses. **GSP**: A simple graphical interface guides users through sequential processing steps. No programming experience is required, but scripts can easily be created for more flexibility. An open ecosystem of plug-ins allows the incorporation of new methods.

FieldTrip (Oostenveld et al., 2011) is developed by a team led by Robert Oostenveld at the Donders Institute (Radboud University, Nijmegen, the Netherlands). This complete MEEG toolbox, also supporting human and animal invasive electrophysiology, is particularly well suited to oscillatory data analysis and source modeling. Its cluster-based permutation statistical analysis framework (Maris and Oostenveld, 2007) is the current gold-standard method to correct for multiple comparisons in MEEG data analysis. A unique feature of this toolbox is its ability to handle trials of varying duration. **GSP**: The systematic separation of configuration parameters and data promotes a robust structure and aids reproducibility. Its command-line interface is suitable for experimenters with basic MATLAB experience, as well as more advanced users.

Brainstorm (Tadel et al., 2011, 2019) was originally created in MATLAB by Baillet et al. (2000) at the Laboratoire de Neurosciences Cognitives et Imagerie Cérébrale (LENA) in Paris. It has been entirely revamped, further developed, and professionally managed by François Tadel since 2008 in Paris, and in Sylvain Baillet’s team at the Montreal Neurological Institute at McGill University in Montreal (Tadel et al., 2011, 2019), with

large contributions from John Moshier (University of Texas Health Science Center at Houston) and Richard Leahy (University of Southern California). Among others, it can handle MEG, fNIRS, scalp, and intracranial EEG within the same protocol. Its sophisticated graphical interface includes an advanced pipeline generator and a modular “process” ecosystem that allows incorporating virtually any analysis method into its environment, including e.g. FieldTrip and MNE-python. GSP: Powerful tools allow data exploration and visualization, and a database engine transparently handles data files.

MNE-Python (Gramfort et al., 2013) is the most recent addition to this toolbox list, based on the older MNE-C Minimum Norm Estimate (MNE) (Gramfort et al., 2014; Hämäläinen and Ilmoniemi, 1994). Alexandre Gramfort (Inria Saclay-Île-de-France Research center, Paris) leads the development. It can handle MEG, EEG, iEEG (ECoG and DBS), EKG, EMG, and fNIRS among others, and interfaces with the Scikit-learn Python toolkit for machine learning analyses (Pedregosa et al., 2011). Its further ecosystem provides powerful analysis and visualization tools, including dipole fitting, linear methods, beamformers, and nonlinear approaches, and interfaces with Freesurfer (Dale and Sereno, 1993; Dale et al., 1999; Fischl, 2012). GSP: The object-oriented interface ensures a match between methods and data types, and a lively contributor community fosters further developments and improvements.

Finally, **SPM** (Statistical Parametric Mapping) was originally developed for the analysis of positron emission tomography (PET) and subsequently fMRI data by Friston et al. (2007). In addition to PET and fMRI, SPM allows analysis of MEEG data (Litvak et al., 2011) using the general linear model (GLM) and a Bayesian framework for source reconstruction. It furthermore supports analyses of effective connectivity and fitting biophysically realistic neural mass models to MEEG data using dynamic causal modeling (DCM; Kiebel et al., 2008). It maintains GSP with support for both a graphical user interface, and scripting analyses for transparent data processing pipelines.

Highlights 2.1.**Human factors**

- Perceptual and cognitive biases influence how researchers perceive, interpret, and communicate data and ideas. Awareness is the first step towards mitigation.
- Careful planning and reasoning can help reduce logical fallacies concerning hypotheses, relevant theories, and conclusions.
- Collective exercises such as a “premortem meeting” can help highlight pitfalls and correct mistakes before they are made during an experimental project.

Highlights 2.2.**Pre-registration**

- Pre-registration is the practice of publicly disclosing and archiving experimental plans before data are collected and/or analyzed.
- Pre-registration procedures can be broadly classified as either “unlocked” (publicly available research plans that might still undergo modifications) or “locked” (finalized research plans, which should be followed precisely throughout the execution of the project).
- Pre-registration has been argued to lead to better scientific practices through increased transparency of the scientific process, peer-review at early stages of the project, and reduced publication bias.
- Potential concerns or limitations such as “scooping”, increased project durations, or reduced applicability for trainees, should be carefully considered but are often mitigated when planning is done carefully.

Highlights 2.3.**Reproducibility and collaboration**

- Tools exist to help prepare experiments in such a way to emphasize reproducibility, replicability, robustness, and generalizability (Box 2).
- Open Science Framework is a discipline-agnostic, general-purpose platform for collaborative, open work.
- The Turing Way is a community-driven guide to reproducible, ethical, inclusive and collaborative data science.
- #EEGManyLabs is a large-scale Initiative for the replication of some of the most influential EEG papers.
- ERP CORE is an open resource to reproduce and replicate common ERP experiments.

Highlights 3.1.**Data collection**

- Careful attention to data collection is crucial as conditions during data collection can affect data quality, with different contexts and different modalities having different requirements.
- It is important to verify the visual and temporal accuracy of the stimulus presentation and its recording, since this is often the main component in an experiment.
- Recording clean data requires specific steps that can be, and have been, collected and published in open acquisition protocols.
- Intracranial measurements provide unique benefits but are limited to clinical settings, where special attention must be given to patient care and other factors.
- Open repositories promote the reproduction of results and data reuse, increasing long-term value of invested efforts.
- Sometimes a project is best served by using simulated data instead of real recordings, e.g., when a known ground truth is needed.

Highlights 3.2.**Analysis software strategies**

- Code is often the only authoritative source of exact experimental procedures, so well-written code greatly aids reproducibility as well as correctness.
- Choosing a software environment to analyze MEEG data requires considering the local lab environment, programming capabilities, and the needs of the experiment being performed.
- Toolbox environments evolve over time, and different implementations of the same procedure can have different outcomes. It is important to document the precise version of any software package or coding environment, and to verify results in different toolboxes when in doubt.

Highlights 3.3.**Signal processing**

- The method used to answer an experimental question sets specific data requirements and should therefore be chosen before designing the experiment.
- Data analysts should understand all signal transformation steps, assumptions, limits and pitfalls of the used methods, as well as the properties of the data itself.
- Running the same analysis with several signal processing alternatives and visualizing the intermediate analysis steps can help assess the robustness of experimental results.

Highlights 3.4.**Statistics**

- Current uses of statistics for univariate and multivariate data-driven approaches require adapting significance thresholds.
- Simulations can help prospective statistical power analyses to aid experimental design.
- Metrics exist to evaluate the data quality of event-related MEEG studies.
- Multiverse analysis can be used to address the issue of the 'garden of forking paths' in MEEG research.

Highlights 3.5.**Machine learning**

- ML algorithms are increasingly popular in MEEG contexts, in particular for data-driven analyses.
- ML algorithms will always provide an outcome, but its relevance and accuracy depend on the quality and meaning of the input data, and may be difficult if not impossible to evaluate neurophysiologically.
- Data preprocessing and experimental design need to take into account known ML pitfalls when ML-based methods are to be used.

Highlights 3.6.**Data standardization, analysis, and reporting**

- Data standardization helps data organization, sharing, and analysis.
- Standards in the form of ‘living documents’ can address ever-evolving issues and perspectives.
- The Brain Imaging Data Structure (BIDS) is a community-led standard for organizing, describing and sharing neuroimaging data.
- ARTEM-IS provides a template and tool for comprehensive reporting.
- COBIDAS MEEG provides guidelines and checklists covering all parts of MEEG projects, specifically including reproducibility and data sharing.

Highlights 4.**Beyond the signal**

- No scientific discipline is free of moral implications, and MEEG research must consider established and proposed rights to e.g. mental integrity.
- Technological innovations and increases in quantity are not goals in and of themselves: questions, hypotheses, and theories remain paramount.
- The culture of speed creates wrong incentives that favor quantity over quality and competition over collaboration, leading to a stressful and uncaring work environment.
- A slow science perspective can help to rethink academia instead of merely fixing it.
- GSP must take environmental aspects into account.

Table 1

Widely used MEEG analysis software toolboxes.

	EGLAB	FieldTrip	Brainstorm	SPM	MNE-python
Toolbox link	https://eeglab.org	https://www.fieldtriptoolbox.org/	https://neuroimage.usc.edu/brainstorm/	https://www.fil.ion.ucl.ac.uk/spm/	https://mne.tools
Environment	MATLAB	MATLAB	MATLAB	MATLAB	Python
Recommended programming level	Beginners to advanced	Intermediate to advanced	Beginners to advanced	Beginners to advanced	Intermediate to advanced
Interface	GUI and scripting	Scripting	GUI and scripting	GUI and scripting	Scripting
Supported modalities	EEG, eye tracking, MoBI	MEG, EEG, iEEG, fNIRS, motion capture, eye tracking, (f)MRI	MEG, EEG, iEEG, fNIRS, Multiunit, motion capture, eye tracking, MRI	(f)MRI, PET, EEG, and MEG	MEG, EEG, iEEG, fNIRS
BIDS support	bids-matlab-tools (Delorme et al., 2021)	data2bids (github.com/fieldtrip/data2bids)	BIDS tools (neuroimage.usc.edu/brainstorm/ExportBids)	en.wikibooks.org/wiki/SPM/BIDS	MNE-BIDS (Appelhoff et al., 2019)
Online community and support	✓	✓	✓	✓	✓
Strengths	<ul style="list-style-type: none"> - Advanced ICA integration - Mobile brain imaging (MoBI) - Advanced statistics including GLM support 	<ul style="list-style-type: none"> - High flexibility - Frequency and time-frequency analysis - Non-parametric statistics 	<ul style="list-style-type: none"> - No programming skills required - Advanced visualization & user interface - Automatic data organization 	<ul style="list-style-type: none"> - Dynamic Causal Modeling (DCM) - Advanced statistics (native GLM support) - Bayesian source analysis framework 	<ul style="list-style-type: none"> - High flexibility - Machine-learning (scikit-learn) - Source estimation
References	Seminal paper Delorme and Makeig (2004), Delorme et al. (2011) Pernet et al. (2021)	Seminal paper Oostenvelde et al. (2011) Popov et al. (2018) Meyer et al. (2021) Andersen (2018)	Seminal paper Tadel et al. (2011)	Seminal paper Friston et al. (1994); Litvak et al. (2011) Henson et al. (2019)	Seminal paper Gramfort et al. (2013), (2014) Jas et al. (2018) Andersen (2018) van Vliet et al. (2018)