ARTEM-IS for ERP: Agreed Reporting Template for EEG Methodology - International Standard for documenting studies on Event-Related Potentials

AUTHORS

Anđela Šoškić[1,2†], Vanja Kovic[2], Johannes Algermissen[3], Nastassja L. Fischer[4], Giorgio Ganis[5,6], Remi Gau[7], Guiomar Niso[8,9], Robert Oostenveld[3,10], Dejan Pajić[11], Mariella Paul[12,13], Yuri G. Pavlov[14,15], Katarina Stekić[2], Yu-Fang Yang[16], Suzy J. Styles[17,4]

[1]Teacher Education Faculty, University of Belgrade, [2]Laboratory for Neurocognition and Applied Cognition, Faculty of Philosophy, University of Belgrade, [3]Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, the Netherlands, [4]Centre for Research and Development in Learning (CRADLE), Nanyang Technological University (NTU), Singapore, [5]School of Psychology, University of Plymouth, UK, [6]Brain Research and Imaging Centre (BRIC), University of Plymouth, UK, [7]Institut des sciences psychologiques, université Catholique de Louvain, Belgium, [8]Psychological & Brain Sciences, Indiana University, Bloomington, USA, [9]Instituto Cajal, CSIC, Madrid, [10]NatMEG, Karolinska Institutet, Stockholm, Sweden, [11]University of Novi Sad, Faculty of Philosophy, Department of Psychology, [12]University of Göttingen, [13]Leibniz Science Campus, Göttingen, Germany, [14]University of Tuebingen, Tuebingen, Germany, [15]Ural Federal University, Ekaterinburg, Russia, [16]Freie Universität Berlin, Berlin, Germany, [17]Psychology, School of Social Sciences, Nanyang Technological University, Singapore

[†]Correspondence should be addressed to Anđela Šoškić; E-mail: andjela.soskic@uf.bg.ac.rs

ABSTRACT (max 300 words)

Given that the choices made during recording, preprocessing and analysis of event-related potentials (ERP) data can affect study outcomes, it is critical that they are transparently reported to allow for reproducibility and replicability. Yet, systematic reviews of reporting practices in the field have shown that journal articles do not meet this goal and that guidelines for writing them better have not resulted in a sufficient improvement to reporting transparency.

ARTEM-IS aims to address this issue by building dynamic, interactive web applications that support documenting information required by existing publication guidelines in the form of a standardised metadata template. Completing an ARTEM-IS form results in a human-reader-friendly PDF and a machine-readable JSON summary of methodological information, which allows for a level of reporting precision higher than what is typically found in journal articles. These can be used as supplements to a publication, as a memory aid when writing a paper, or as records that allow easier metadata extraction in comparison to verbal descriptions in papers.

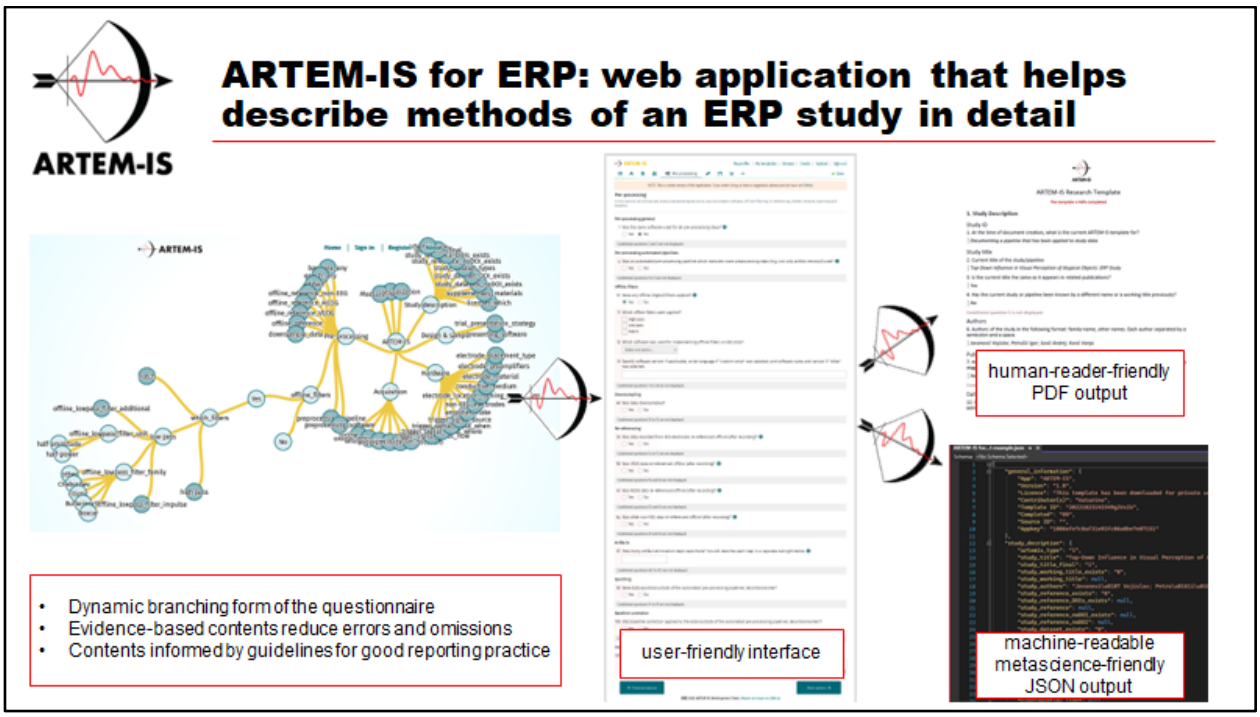
Here, we present the ARTEM-IS for ERP, which supports describing a typical ERP study, including most of its core methodological aspects (study description, experimental design,

hardware, data acquisition, pre-processing, measurement, visualisation, additional comments). We discuss the current contents of the form, web application functionalities, current limitations, and potential directions for future developments. In addition, the process of building the form contents and the web application through a collaborative grassroots initiative is described. Finally, we argue that a wider adoption of ARTEM-IS can bring benefits to different stakeholders: researchers themselves or their collaborators, especially on large-scale projects, reviewers, readers of a paper, and the scientific community at large.

KEY WORDS (max 5)

Event-related potentials, transparency, standardisation, open science, ARTEM-IS

GRAPHICAL ABSTRACT



HIGHLIGHTS

- Transparent reporting of ERP methods is needed for reproducibility and replicability
- Journal article methods sections are not sufficiently detailed for this goal
- ARTEM-IS for ERP is a web application that helps describe ERP methods in detail
- Completing the form results in a human-reader-friendly and a machine-readable output
- They can serve as paper supplements, memory aid when writing, or metadata records

ABBREVIATIONS

ARTEM-IS - Agreed Reporting Template for EEG Methodology - International Standard
BIDS - Brain Imaging Data Structure

COBIDAS - OHBM Committee on Best Practice in Data Analysis and Sharing
CSV - Comma Separated Values
EEG - electroencephalography
EOG - electrooculography
ERP - event-related potentials
ERN - error-related negativity
ICA - independent component analysis
INCF - International Neuroinformatics Coordinating Facility
JSON - JavaScript Object Notation
MEEG - magneto- and electroencephalography
MEG - magnetoencephalography
MySQL - My Structured Query Language
OHBM - Organisation for Human Brain Mapping
OSF - Open Science Foundation
PDF - Portable Document Format
PHP - PHP: Hypertext Preprocessor

ACKNOWLEDGEMENTS

D.P. was supported by grant 7744418, Genetic and environmental influences on psychological adaptation of children and adults – GENIUS (Science Fund of the Republic of Serbia); A.Š. was supported by 451-03-1/2022-14/4 (Ministry of Education, Science and Technological Development of the Republic of Serbia); V.K. and K.S. were supported by grant 451-03-9/2021-14/ 200163 (Ministry of Education, Science and Technological Development of the Republic of Serbia); N.L.F. was supported by grant Cambridge-NTU Centre for Lifelong Learning and Individualised Cognition - CLIC (National Research Foundation (NRF), Prime Minister's Office, Singapore, under its Campus for Research Excellence and Technological Enterprise (CREATE) Programme); Y.G.P. was supported by grant PA 4005/1-1 (German Research Society (Deutsche Forschungsgemeinschaft, DFG)); S.J.S. was supported by grant NRF2016-SOL002-011 (Singapore's National Research Foundation under the Science of Learning CRP).

The authors would like to thank contributors at hackathons: Alec Shaw, Ekansh Sareen, Jasmine Wei Rou Tan, Shamala Sundaray, Defu Yap, Fei Ting Woon, and Han Ke, and extend their thanks to Jasmine Wei Rou Tan and Han Ke for taking part in the activities of the INCF Working Group on ARTEM-IS.

AUTHOR NOTE

Credit roles:
Conceptualization (originated or substantially developed the ARTEM-IS Project or ways of working)**:** Anđela Šoškić, Vanja Kovic, and Suzy J. Styles.
Data curation (curation of OSF or GitHub, creation of shared resources, or Zotero library)**:** Anđela Šoškić, Vanja Kovic, Remi Gau, Yu-Fang Yang, and Suzy J. Styles.

Methodology (creating or adapting data fields in tables of the ARTEM-IS templates or related assets)**:** Anđela Šoškić, Vanja Kovic, Johannes Algermissen, Nastassja L. Fischer, Giorgio Ganis, Remi Gau, Guiomar Niso, Robert Oostenveld, Dejan Pajić, Yuri G. Pavlov, Katarina Stekić, Yu-Fang Yang, and Suzy J. Styles.
Project administration (planning meetings, hacks, chairing/hosting group events and organising cross-group meetings)**:** Anđela Šoškić, Vanja Kovic, and Suzy J. Styles.
Software (programming of any kind, e.g., supporting the conversion of base metadata tables into GUIs)**:** Remi Gau, Dejan Pajić, and Katarina Stekić.
Validation (checking user interfaces OR cross-checking tabulated data of any kind)**:** Anđela Šoškić, Vanja Kovic, Johannes Algermissen, Nastassja L. Fischer, Giorgio Ganis, Guiomar Niso, Robert Oostenveld, Yuri G. Pavlov, Katarina Stekić, Yu-Fang Yang, and Suzy J. Styles.
Visualisation: Anđela Šoškić, Vanja Kovic, Giorgio Ganis, Dejan Pajić, and Suzy J. Styles.
Writing - original draft: Anđela Šoškić, Vanja Kovic, Nastassja L. Fischer, Giorgio Ganis, Guiomar Niso, Robert Oostenveld, Mariella Paul, Yuri G. Pavlov, Katarina Stekić, Yu-Fang Yang, and Suzy J. Styles.
Writing - review & editing: Anđela Šoškić, Vanja Kovic, Johannes Algermissen, Nastassja L. Fischer, Giorgio Ganis, Robert Oostenveld, Dejan Pajić, Mariella Paul, Katarina Stekić, Yu-Fang Yang, and Suzy J. Styles.

The order of authors is alphabetical except for the first, the second and the last position. All authors have read and approved the final version of the text.

1. Introduction

Recordings of brain activity and behaviour in neuroscientific experiments result in very rich datasets. The high spatial and temporal density of these datasets allows for a myriad of possible data processing and analysis pathways, sometimes known as the ‘garden of forking paths’ (Gelman & Loken, 2013) as every analytical decision contributes to an ever increasing number of possible outcomes (paths through the garden). Specifically, this problem is present in electroencephalography (EEG) research and, within EEG, in event-related potential (ERP) research. In order to go from raw EEG data to ERPs, the focus of this paper, the data has to undergo a complex yet flexible sequence of preprocessing and analysis steps (referred to as a “pipeline”). For each of these steps, researchers have to make decisions, including how to filter and reference the raw data, which baseline correction to apply, and which statistical analysis to apply.

Recent studies have shown that these decisions matter. Sandre et al. (2020) applied several different common implementations of some preprocessing steps to the same ERP dataset and found that the preprocessing choices affected the effects on the ERN component as well as its reliability. In another recent demonstration in the domain of the N400 family of ERP components, one chain of decisions in which 8 steps were varied with only two to four decisions per step resulted in 864 possible pre-processing and analysis pipelines, and the significance and size of

some of the experimental effects were shown to fluctuate across 14 chosen pipelines (Šoškić et al., 2022). Furthermore, the selection of software toolboxes or versions may contribute to variations in the results (Kabbara et al., 2022; Lefebvre et al., 2018). Choices in ERP methodology are thus very likely to be one explaining factor for inconsistencies in the literature and differences between original studies and replications.

Given that the choices made during preprocessing and analysis of ERP data can have an effect on the results, it is critical that all choices are transparently reported to allow for reproducibility and replicability (Niso et al., 2022). With the hope of facilitating transparent and complete reporting of methodological details, several ERP reporting guidelines have been published (Donchin et al., 1977; Keil et al., 2014; C. Pernet et al., 2020; Picton et al., 2000). The trend towards transparency in ERP and EEG research is part of a broader effort to increase transparency and reproducibility of the scientific record in the past two decades (e.g., Foster & Deardorff, 2017; Larson & Moser, 2017; Niso et al., 2022; Open Science Collaboration, 2015; Wilkinson et al., 2016).

The requirement for transparent and comprehensive reporting of preprocessing and analysis details in neuroscience is also critical from an ethical perspective, especially for clinical research (Emanuel et al., 2000). The ethical principle of beneficence explicitly requires researchers to minimise the risk/benefit ratio for participants in their studies (e.g., National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979). It is hard to ensure an ethical balance when scientific development is slowed down by contradictory findings and uncertainties in the literature. For example, when there is limited or incorrect reporting of methodological preprocessing and analysis details, then human studies need to be carried out multiple times in different laboratories to resolve contradictions in the literature, resulting in undue burden on participants (see Van Dang, 2020) for an example of this in cancer research).

However, the available evidence indicates that reporting guidelines and calls for increasing transparency have not had an impact on the amount of methodological detail provided by traditional journal articles in the domain of ERP research (Clayson et al., 2019; Šoškić, 2021; see also Paul & Mani, 2022), suggesting that new approaches may be needed.

2. Existing efforts towards tools for methodology documentation

A pioneering effort towards improving the reporting of ERP studies through novel approaches was the MEG and EEG (MEEG) publication guidelines by Keil et al. (2014), which featured a checklist/reminder of items to include when writing an MEEG paper, including the ones on ERP studies. The list contained 51 items focusing on different research steps, from hypotheses to figures and statistical analysis details. Even though this list provided useful instructions, research has shown that it has not had an immediate impact on the quality of documenting methodology in ERP papers before and after the publication of the checklist (Clayson et al., 2019; Šoškić,

2021). One possible explanation is that pressure to publish fast in traditional journal articles with limited space can outweigh the perceived value of an additional checklist (Styles, Ković, et al., 2021).

Another initiative has been launched by the Committee on Best Practices in Data Analysis and Sharing (COBIDAS) by the Organisation for Human Brain Mapping (OHBM), which is developing guidelines for good practice for MRI (Nichols et al., 2016)and MEEG (Pernet et al., 2018). These guidelines include checklists that are more detailed than the checklist offered by Keil et al (2014). Noting the ever-increasing complexity of analysis pipelines, Pernet et al. (2018) recognised that the need to describe pipelines thoroughly is in conflict with limited journal space, and they recommended preparing supplementary tables with details listed in their checklist. To make this easier, COBIDAS and a working group from the International Neuroinformatics Coordinating Facility (INCF) are working on a web-app (eCOBIDAS) that would allow easier creation of such supplements (Gau et al., 2022; https://www.incf.org/sig/incfohbm-working-group-checklists-transparent-methods-reporting-neuroscience-ecobidas). Unlike the checklist by Keil et al. (2014), which is meant to be completed at the time of writing a manuscript, this tool separates the process of methods documenting from writing the paper, making it possible to integrate the documentation process with all other research stages (e.g., decisions can be documented during planning a study or a change can be documented when a new decision is made during data collection or analysis). This feature is known to provide benefits to checklist completion (Degani & Wiener, 1993; see Styles et al., 2021, for a discussion of this question in the context of EEG specifically).

In addition to projects specifically focused on improving methodological descriptions in papers, other initiatives that aim at facilitating research transparency include an ERP pre-registration template (Govaart et al., 2022), efforts aimed at standardising rather than documenting pipelines (e.g., Bigdely-Shamlo et al., 2015; Gabard-Durnam et al., 2018; Monachino et al., 2021; Rodrigues et al., 2021), and data/metadata sharing standards and platforms (e.g., BIDS - Gorgolewski et al., 2016; Pernet et al., 2019) and the BIDS-compliant data sharing platform OpenNeuro, EEGBase portal (Papež & Mouček, 2017). The pre-registration template can help researchers consider, make and document methodological decisions before conducting a study. However, although pre-registrations are a useful reminder of the study's decision-making process, they are not aimed at documenting what eventually happened when the research was conducted. Similarly, standardised data pre-processing pipelines are not study documentation tools, but they have an additional benefit of being easier to describe and thus have potential for improved reproducibility compared to idiosyncratic pipelines, as pointed out by researchers developing these pipelines (e.g., Rodrigues et al., 2021). Finally, data standardisation and sharing platforms can help preserve very detailed methodological information in a metadata file accompanying a dataset, including information even on, for example, individual participants or trials. However, creating well-curated metadata describing complex pipelines, typically conducted using multiple software programs, can be demanding, and some of the methodological decisions, especially the rationale in decision-making, are still not preserved in metadata, data itself or code. In addition, extracting all desired information from metadata stored in this form is time-consuming for reviewers and other researchers.

3. ARTEM-IS as a methodology documentation initiative

Taking into account the challenges described above, ARTEM-IS (short for Agreed Reporting Template for EEG Methodology - International Standard) aims to provide tools that make documenting methodology easier and more transparent, and thus improve reproducibility of research.

More specifically, the goal of ARTEM-IS is to create interactive tools for methods reporting in the form of web applications which lead users through a set of questions from study design to data visualisation, including the rationale for decisions made along the way. To make the process intuitive and foolproof from omissions and errors, ARTEM-IS forms[1] are simple to use and require answers as short and as specific as possible. Completing the form results in a summary of methodological information that achieves a level of compliance with recommendations for good reporting practices much higher than what is typically found in journal articles. Outputs of the ARTEM-IS web application can be stored online and downloaded as human-reader-friendly PDF and/or as machine-readable JSON documents. These can be used as supplements to a publication, as a memory aid when writing a paper, or as records that allow easier metadata extraction in comparison to verbal descriptions in papers.

Rather than using yes/no checklists of whether a detail has been reported, the users are given a form comprised of individual data fields to be filled in. This allows both (1) documenting the decisions as they are being made, as discussed in the previous section, and (2) reducing omissions that arise when an incomplete or skipped item is mistakenly marked as complete in a yes/no checklist (Styles, Kovic, et al., 2021). Documenting methodological decisions at any stage of the research process allows ARTEM-IS to serve for documenting pre-registration plans, studies in progress, as well as completed studies. Using the template prior to conducting a study can be especially helpful to researchers who are new to EEG, but it can serve as a useful reminder for seasoned researchers.

Documenting a study in detail sufficient for replication attempts and for metascience purposes may sound time consuming at first. However, unlike traditional guidelines and checklists, ARTEM-IS does not show a static and exhaustive list of steps researchers might potentially consider. Instead, the online web application has a dynamic branching structure that mimics the decision-making process itself by presenting questions contingent on previous answers and thus allows skipping paths not taken, reducing the workload considerably. Moreover, most questions require selecting predefined options or inputting short answers. This not only facilitates completing the form, but also comparison between different studies. Third, once filled in, information on features

[1] In this paper, when we refer to the ARTEM-IS *form*, we have in mind the online questionnaire that is filled in by users in the ARTEM-IS web application. ARTEM-IS *template* is used to refer to the standardised format ("mould") which is defined by questions in the form and following which all outputs from the web-app are structured in the same way. When we refer to an ARTEM-IS *tool*, we have in mind the web application with all its functionalities, together with the structure and contents of the ARTEM-IS form.

such as EEG equipment can be used as a template for future study descriptions, shortening the process of study documenting even further. In addition, the time invested in detailed documentation can have multiple investment returns in the long run for different stakeholders: researchers themselves or their collaborators, especially on large-scale projects, reviewers, readers of a paper, the scientific community at large. Examples of these returns include more accurate pre-registrations; more and easier replication attempts, meta-analytic projects, and large-scale collaborations; easier evaluation of reports by reviewers, editors, and readers (possibly even in a semi-automated manner in the future); reduction of errors in descriptions of methods in the literature; less need for requests for additional information from corresponding authors; and a better understanding of the sources of inconsistencies in the literature.

Finally, ARTEM-IS tools allow documenting a study to a higher level of precision than any existing tools by (1) adhering to ARTEM-IS Design Guidelines, which have been developed by leveraging expertise of other fields with a longer history of using checklists, such as surgery and aviation, as well as by (2) utilising insights into reporting errors and omissions common in the EEG literature, which are obtained using detailed systematic reviews of methodology reporting (e.g., Šoškić, 2021). The general ARTEM-IS approach is discussed in more detail by (Styles, Ković, et al., 2021).

4. ARTEM-IS for ERP: the first ARTEM-IS documenting tool

The main topic of the present paper is to describe the first ARTEM-IS tool: ARTEM-IS for ERP, designed for documenting a study on event-related potentials. The up-to-date version of the template can be accessed at http://artemis.incf.org/ and is maintained at https://github.com/INCF/artem-is. In the following sections, we will describe its contents, how it was developed and its limitations.

4.1. Form structure

In its current state (version 1.0), ARTEM-IS for ERP comprises nine sections designed to assist researchers report the following aspects of a study: (1) general description, (2) experimental design, (3) hardware, (4) acquisition, (5) pre-processing, (6) channel selection for analysis, (7) measurement, (8) visualisation, and (9) other. The description of statistical analysis is not included in this version. Each section contains a set of questions with answers that can be categorical, numerical or short open-ended statements:

1. **Study**: This section gathers information about general aspects of the research study: information related to whether this is a report to document a planned pipeline or for an already applied one; the title of the study; information related to the authors; related

publications and DOIs; associated available datasets and supplementary materials (e.g., code); licensing; abstract; keywords; financial support; ethics committee approval; acknowledgements; and citation instructions.

2. **Experimental design**: This section gathers information about the experimental design and the sample of the study: the number of participants in the study; inclusion and exclusion criteria; information related to the experimental/comparison groups; details on the trials presented and analysed; and software for stimuli presentation.
3. **Hardware**: This section provides information about the hardware used for recording and acquisition: information related to the EEG cap/net and electrodes characteristics and electrode placement scheme; information on the amplifier characteristics; description of triggers (e.g., how they were generated and saved); and information related to any additional devices used (e.g., signal boxes, converters, electrode position measurement devices).
4. **Acquisition**: This section documents information related to the data acquisition process: acquisition software used and version; details on impedances or alternative data acquisition quality measures; references for EEG, EOG and other electrodes, if used, and their alternatives (e.g., Driven Right Leg (DRL)); ground electrode and placement; EOG channels for recording eye movement-dependent voltage and placement; online high pass and low pass filters and notch filters; and acquisition sampling rate.
5. **Pre-processing**: This section documents pre-processing, i.e., the steps in the pipeline before measurement and statistical analysis of an ERP component: automated preprocessing pipeline; offline filtering; downsampling; re-referencing of EEG, EOG and other channels, if used; artifact removal methods (rejection of bad trials, data segments, or channels, artifact correction, channel interpolation, multi-step automated approaches); epoching, baseline correction. For each of these steps, the user can indicate whether it was applied, which software was used if it is indicated that the software was not the same in all pre-processing stages, and they can describe appropriate settings. At the end of this section, the user is asked to specify the order in which pre-processing steps were applied.
6. **Measurements**: This section offers options to describe amplitude and latency measurements of one ERP component, unless the user subjected all time points and channels to statistical analysis (e.g., in a mass univariate analysis). Amplitude properties include: measurement software, amplitude measure (peak, mean, window area, etc.) and its parameters, waveforms used to measure amplitude, measurement time window, rationale of selecting this exact time window. Similarly, latency properties include: measurement software, midpoint latency measure (e.g., local peak, 50% area) - if applied, onset latency measure (e.g., fractional area, fractional peak) - if applied, the appropriate parameters of each latency measure depending on the choices, time window within which the latency was searched for, and the rationale for selecting this exact time window.
7. **Channels**: This section gathers information about channel selection for later statistical analysis. Like in the case of the previous section, the current version of the template supports describing the location for the measurement of one ERP component. There are two main scenarios - that all electrodes are included in statistical analyses (e.g., in a mass univariate approach), and that a subset of electrodes is included in the analyses. Within the second scenario, five options are offered: a priori selection of channels; data-driven

selection of channels from the entire scalp; data-driven selection of channels from an a priori selected broader region on the scalp; selection/identification of channels in two data driven steps (e.g., maximal effect within a visually identified broader region); other. Depending on the choice, an appropriate subset of questions follows. In the end of this section, the user is asked to describe whether and how the channels used for analysis were aggregated into regions of interest.

8. **Visualisation**: This section provides information related to the visualisation: the type of plot created (e.g., line plot (waveforms), topoplot); if any extra pre-processing was performed only for visualisation purposes (e.g. smoothing filter, different baseline); description of what the data represents (e.g. single waves, difference waves) and corresponding units (e.g. voltage maps, normalised voltage maps); description of which conditions or difference waves were shown; channels selected for visualisation and the rationale for this decision, beginning and end of the time window visualised and subdivision of this time window into shorter segments for topographical plots.
9. **Other**: This section contains an option to add additional comments, if there are any.

A quick overview of all the items of the checklist, without any branching logic, can be found on the ARTEM-IS OSF project (https://osf.io/z3tjb). A dynamic overview of the branching structure of multiple-choice questions in the form is available on the front page of the web application (http://artemis.incf.org/), where it is visible before logging in, while an example of this branching can be seen in Figure 1.

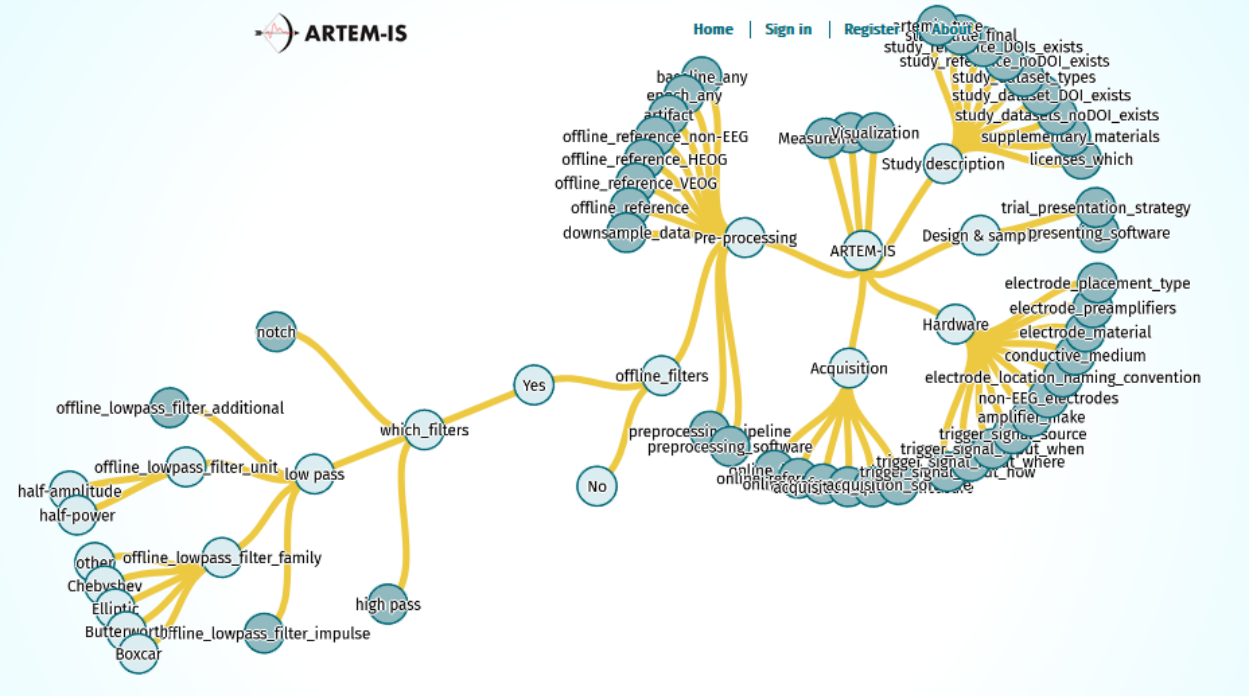


Figure 1. Part of the branching structure of ARTEM-IS for ERP reporting items. Nodes at the first level of branching represent sections, while other nodes represent either multiple-choice items (questions) or choices that can be picked.

4.2. Web application description

Existing off-the-shelf software solutions were not suitable for providing a flexible and fully controlled environment to serve as a platform on which ARTEM-IS would be built. Hence, the ARTEM-IS team has opted for the development of a custom web application as an interface for the dynamic and interactive data framework. Developing an application from scratch ensured the future upgrading and improvement of the ARTEM-IS for ERP, as well as building ARTEM-IS tools for other subdomains of EEG research, by providing an easy way to modify and adapt database structure and thus the structure of the general form itself.

The first motivation behind the creation of the application was to enable users to easily generate ERP research reports using the guided interactive form. Questions are organised into sections, so that one can switch between them, not necessarily filling the questionnaire in a linear way. Changes to the template are saved automatically, thus making it possible to document incomplete templates for research in progress. Although there are no mandatory questions, users are provided with feedback on the percentage of template completion, i.e., the percentage of answered questions. This information is created dynamically since the number of such items in the questionnaire may vary. Most of the questions are associated with a logical expression that determines whether the question should be displayed or not, depending on the answers to previous questions. Users always have feedback on how many questions still require response since they are clearly marked in red.

The second motivation behind the development of the ARTEM-IS application was to facilitate the exchange of information and collaboration among researchers. The application is free to use but requires user registration since the user is responsible for the decision on who would be able to see, download, and/or modify the template and under what conditions. Reports can be restricted to an invited group of contributors or distributed freely under the CC BY 4.0 licence. In this way, the creator of the report decides whether their research team should jointly work on a closed document with restricted access, or the report would be available for all registered users to copy and adapt it for the purposes of their own research. Contributors to the report have the option to add new or existing ARTEM-IS users to the team and invite them to collaborate on the project report. New collaborators automatically receive an email notification with the unique link to the ARTEM-IS template form. Information on authors of a study is separate from the information on the contributors to its ARTEM-IS report, allowing, for example, researchers to systematically document the method of papers that are published by others for the purposes of systematic reviews and meta-analyses. As mentioned earlier, the ARTEM-IS application also enables users to export research details both in the form of a PDF report and JSON interchange file. Structured JSON objects, unlike, for example, CSV files, provide a convenient way to store hierarchically organised, machine-readable information as a set of key/value, i.e., question/response pairs. Additionally, JSON files are exported in the so-called “pretty-print” format which enables users to easily search and browse visually through the template structure. An example of the output in both formats can be found on the OSF page of this project (https://osf.io/ahp3t/). JSON format

provides a convenient way for users to export and import information and build their own research report upon the already existing privately shared or open access templates.

ARTEM-IS questionnaire structure and contents were defined in a Google spreadsheet to facilitate collaborative work with a low bar of entry for new contributors in terms of technical skills (see the following section on community involvement). The Google spreadsheet contains separate sheets for each section, and each sheet has columns that describe questions, which are organised in rows. The columns include properties such as the metadata field name in the JSON file for each question, question text, logical expression defining the conditions under which the question is shown, type of input that is expected, choices that are offered for multiple choice questions, BIDS keywords corresponding to a question to facilitate conversions in the future (this part is still in progress). A static version of this spreadsheet (https://osf.io/9dweq/) and its brief description (https://osf.io/ahp3t/wiki/home/) can be found on the OSF page of this project. This spreadsheet is compatible with the format used by the INCF eCOBIDAS Working Group, which allows for compatibility with eCOBIDAS web applications and consequently, easier collaboration on shared goals. The web application was created in PHP and JavaScript, using MySQL as a database backend.

Besides the web application and the Google spreadsheet, the content of the ARTEM-IS for ERP currently also exists in two additional formats: (1) as a simplified “linear” spreadsheet for a quick overview of all the items of the checklist without any branching logic (https://osf.io/z3tjb); and (2) as a set of JSON linked data files relying on the Reproschema (Reproschema Contributors, 2021) to provide a representation of the whole ARTEM-IS template that is both standardised according to a schema and independent of the technological decisions made for its implementation and rendering as a web-app (https://github.com/INCF/artem-is/tree/main/schemas/artemis/activities).

4.3. Development through community involvement

ARTEM-IS for ERP has been developed as a collaborative grassroots initiative which has involved contributors from different backgrounds who were recruited at conferences, hackathons and via social media calls, and it is open for further improvements through scientific community collaboration. In line with the ARTEM-IS Guidelines, ARTEM-IS for ERP has been initially developed on the basis of the outcomes of the systematic review of ERP literature by Šoškić et al. (2021). Since then, it has been continually improved through efforts of the INCF working group on ARTEM-IS (https://www.incf.org/sig/incf-working-group-artem) with contributions from a wider group of collaborators at hackathons (Gau et al., 2022), specifically OHBM BrainHack in July 2021 (https://github.com/ohbm/hackathon2021/issues/11) and BrainHack Global in December 2021 (https://brainhack.org/global2021/project/project_153/), and with beneficial exchanges with other projects on open neuroscience, in particular eCOBIDAS. Full list of contributors is maintained on OSF (https://osf.io/ut9pc/). In addition, community feedback was requested through consultation opportunities at conferences, such as LiveMEEG 2020 (Šoškić et al., 2020), PuG (Šoškić, 2021), SIPS (Yang & Styles, 2022), INCF Assembly (Šoškić et al., 2022), SPR (Šoškić, Ković, et al., 2022b) or Neuronus (Šoškić, Ković, et al., 2022a).

The benefits of community involvement become obvious when comparing the original draft of ARTEM-IS for ERP to the current version. The original draft comprised a static spreadsheet table with 92 fields to be filled in, with options to add more detail next to each field. The current v1.0 version of ARTEM-IS for ERP is a web application with 337 fields that facilitates dynamic and structured data entry and that has all features described above.

4.4. Limitations of ARTEM-IS v1.0

The ARTEM-IS for ERP v1.0 is exactly that, the first version of the template that is fully functional and can be used to document an ERP study. As the first version, it includes several core features that can be used and tested by a broader community of users, but that can be further built upon in the next versions.

The web-app currently allows documenting a simple ERP experiment involving one ERP component (one set of electrodes; latency and amplitude measurements on this set of electrodes and each from one time window; for documenting another ERP component, users could simply export and import the JSON output, and change the parameters that are different for the new component). It currently does not cover the statistical analysis, which is itself so complex and varied that it could easily become a separate ARTEM-IS template. Until such a time, users may choose to report their statistical analysis in the “Other” section if there is need. The web-app design allows expanding the scope of the web-app to include more complex designs and statistical analysis in the future.

In addition, in the current version of the form, some steps (e.g., artifact removal decisions that involve more complex algorithms and procedures, such as ICA) are still left to the user to describe them in an open-ended format. These can be expanded into a full branching structure of the shortest possible questions to be fully compliant with ARTEM-IS Guidelines in subsequent versions.

Finally, there is space for further enhancements and refinements as ERP methodology progresses and as we learn more from the community using the tool. For example, in future editions, new options may be added to categorical questions that now fit under the category of “Other”, or new questions may be added to include new developments in science or to break down existing open-ended questions. On the other hand, the existing tool can be made more accessible and educative for researchers new to ERP by preparing instructional materials, such as a lexicon of used terminology, help files and detailed elaboration-and-extension documents. One stretch goal of the project is to consider ARTEM-IS as a pathway to software that suggests automatically generated text suitable for use in a methods section.

5. General Conclusions

ARTEM-IS addresses the emerging needs of scholarly publications in the field of electrophysiology by offering EEG methodology metadata templates, which help researchers document methodological decisions from study design to data visualisation in a reproducible and metascience-friendly way. To ensure both quality of the tools and equity of access to them, ARTEM-IS templates are developed with the input from the wider EEG community and openly available to everyone. ARTEM-IS for ERP is the first ARTEM-IS tool, focused specifically on documenting studies using event-related potentials (http://artemis.incf.org/).

Documenting an experiment to the level of detail sufficient for replication attempts and metascience has several benefits. For example, the ARTEM-IS for ERP could potentially fast forward ERP reporting, as it can help researchers prepare the Methods section of manuscripts more easily and accurately. Not only reporting, but the review process could also be made easier, if journal editors ask researchers to attach the ARTEM-IS template with article submission. This would allow reviewers to check more easily whether crucial steps in the reporting pipeline have been skipped, reported ambiguously, or performed in a manner inconsistent with other elements of the methodological description. In addition, improved reproducibility and replicability are in line with the ethical principle of beneficence, given that low reproducibility and replicability slow down the progress of science and put undue additional burden on participants. Another advantage of using ARTEM-IS for ERP is that it allows researchers to outsource some methodology details to the supplementary template, saving more of the valuable journal space for results and discussion. Next, while guidelines for pre-registering ERP studies exist (Paul & Mani, 2022), ARTEM-IS for ERP can make *pre-registration* efforts more accurate due to the high level of precision it enables. Documenting study methodology before data collection is not only helpful for improving transparency, but also for practising careful planning by reducing unnoticed flexibility. Similarly, when *documenting completed studies*, promoting good scientific practice (Niso et al., 2022) of sharing a detailed analysis pipeline will benefit both researchers interested in replications and the authors themselves when designing new studies based on previous findings (Garrett-Ruffin et al., 2021). Finally, leveraging data from the ARTEM-IS reporting template could inspire future *replication* studies and promote open science. Namely, there have not been many replication efforts in the field of EEG so far (Pavlov et al., 2021; Šoškić, 2021), and one of the reasons may be that it is difficult to design replications without deep involvement of original authors, as initial steps in the #EEGManyLabs project has shown (Pavlov et al., 2021). Moreover, helping researchers achieve closer replications would allow opening up for discussions around replication results to be centred around theoretical questions rather than methodological differences.

In the current version, ARTEM-IS for ERP v1.0 includes most of the core aspects of methods of a typical ERP study, and as such it is already ready to use for a broad class of typical ERP studies. While we believe that this tool is already a big step forward in documenting methodological decisions in the field of ERP, it can be further improved by adding more sections and questions, primarily to allow documenting more complex designs and statistical analysis, but also by keeping up to date with progress and trends in ERP methodology. As with all previous developments, this

work in progress is also open to input from the entire ERP community through contributions and feedback at future conferences and hackathons or joining the work of the INCF Working Group on ARTEM-IS.

References

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*, 16. https://doi.org/10.3389/fninf.2015.00016

Clayson, P. E., Carbine, K. A., Baldwin, S. A., & Larson, M. J. (2019). Methodological reporting behavior, sample sizes, and statistical power in studies of event-related potentials: Barriers to reproducibility and replicability. *Psychophysiology*, *56*(11), e13437. https://doi.org/10.1111/psyp.13437

Degani, A., & Wiener, E. L. (1993). Cockpit checklists: Concepts, design, and use. *Human Factors*, *35*(2), 345–359. https://doi.org/10.1177/001872089303500209

Donchin, E., Callaway, E., Cooper, R., Desmedt, J. E., Goff, W. R., Hillyard, S. A., & 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* (pp. 1–11). Karger.

Emanuel, E. J., Wendler, D., & Grady, C. (2000). What makes clinical research ethical? *JAMA : The Journal of the American Medical Association*, *283*(20), 2701–2711. https://doi.org/10.1001/jama.283.20.2701

Foster, E. D., & Deardorff, A. (2017). Open science framework (OSF). *Journal of the Medical Library Association: JMLA*, *105*(2), 203. https://doi.org/10.5195/jmla.2017.88

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

Garrett-Ruffin, S., Hindash, A. C., Kaczkurkin, A. N., Mears, R. P., Morales, S., Paul, K., Pavlov, Y. G., & Keil, A. (2021). Open science in psychophysiology: An overview of challenges and emerging solutions. *International Journal of Psychophysiology*, *162*, 69–78. https://doi.org/10.1016/j.ijpsycho.2021.02.005

Gau, R., Gould van Praag, C., van Mourik, T., Wiebels, K., Adolfi, F., Scarpazza, C., Ruotsalainen, I., Tepper, A., Sjoerds, Z., Simon, J., Klapwijk, E., Hortensius, R., Bartlett, J., & Moreau, D. (2022). *COBIDAS checklist*. https://doi.org/10.17605/OSF.IO/ANVQY

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. *Department of Statistics, Columbia University*, *348*, 1–17.

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., … Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*(1), 160044. https://doi.org/10.1038/sdata.2016.44

Govaart, G. H., Schettino, A., Helbling, S., Mehler, D. M. A., Ngiam, W. X. Q., Moreau, D., Chiossi, F., Zanesco, A. P., Yang, Y.-F., Gau, R., Bartlett, J. E., García Alanis, J. C., Gutsell, J. N., Çetinçelik, M., Pavlov, Y. G., Šoškić, A., Ehinger, B. V., Mouseli, P., Algermissen, J., … Paul, M. (2022). *EEG ERP Preregistration Template* [Preprint]. MetaArXiv. https://doi.org/10.31222/osf.io/4nvpt

Kabbara, A., Forde, N., Maumet, C., & Hassan, M. (2022). *Successful reproduction of a large EEG study across software packages* [Preprint]. Neuroscience. https://doi.org/10.1101/2022.08.03.502683

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., Luu, P., Miller, G. A., & Yee, C. M. (2014). Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology*, *51*(1), 1–21. https://doi.org/10.1111/psyp.12147

Larson, M. J., & Moser, J. S. (2017). Rigor and replication: Toward improved best practices in human electrophysiology research. *International Journal of Psychophysiology*, *111*, 1–4. https://doi.org/10.1016/j.ijpsycho.2016.12.001

Lefebvre, A., Delorme, R., Delanoë, C., Amsellem, F., Beggiato, A., Germanaud, D., Bourgeron, T., Toro, R., & Dumas, G. (2018). Alpha Waves as a Neuromarker of Autism Spectrum Disorder: The Challenge of Reproducibility and Heterogeneity. *Frontiers in Neuroscience*, *12*, 662. https://doi.org/10.3389/fnins.2018.00662

Monachino, A. D., Lopez, K. L., Pierce, L. J., & Gabard-Durnam, L. J. (2021). *The HAPPE plus Event-Related (HAPPE+ER) Software: A Standardized Processing Pipeline for Event-Related Potential Analyses* [Preprint]. Neuroscience. https://doi.org/10.1101/2021.07.02.450946

Nichols, T. E., Das, S., Eickhoff, S. B., Evans, A. C., Glatard, T., Hanke, M., Kriegeskorte, N., Milham, M. P., Poldrack, R. A., Poline, J.-B., Proal, E., Thirion, B., Van Essen, D. C., White, T., & Yeo, B. T. T. (2016). *Best Practices in Data Analysis and Sharing in Neuroimaging using MRI* [Preprint]. Neuroscience. https://doi.org/10.1101/054262

Niso, G., Krol, L. R., Combrisson, E., Dubarry, A.-S., Elliott, M. A., François, C., Héjja-Brichard, Y., Herbst, S. K., Jerbi, K., Kovic, V., & others. (2022). Good scientific practice in MEEG research: Progress and perspectives. *NeuroImage*, 119056. https://doi.org/10.1016/j.neuroimage.2022.119056

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

Papež, V., & Mouček, R. (2017). Applying an Archetype-Based Approach to Electroencephalography/Event-Related Potential Experiments in the EEGBase Resource. *Frontiers in Neuroinformatics*, *11*. https://doi.org/10.3389/fninf.2017.00024

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

Pavlov, Y. G., Adamian, N., Appelhoff, S., Arvaneh, M., Benwell, C. S. Y., Beste, C., Bland, A. R., Bradford, D. E., Bublatzky, F., Busch, N. A., Clayson, P. E., Cruse, D., Czeszumski, A., Dreber, A., Dumas, G., Ehinger, B., Ganis, G., He, X., Hinojosa, J. A., … Mushtaq, F. (2021). #EEGManyLabs: Investigating the replicability of influential EEG experiments. *Cortex*, *144*, 213–229. https://doi.org/10.1016/j.cortex.2021.03.013

Pernet, C., Garrido, M. I., Gramfort, A., Maurits, N., Michel, C. M., Pang, E., Salmelin, R., Schoffelen, J. M., Valdes-Sosa, P. A., & Puce, A. (2020). Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research. *Nature Neuroscience*, *23*(12), 1473–1483. https://doi.org/10.1038/s41593-020-00709-0

Pernet, C. R., Appelhoff, S., Gorgolewski, K. J., Flandin, G., Phillips, C., Delorme, A., & Oostenveld, R. (2019). EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Scientific Data*, *6*(1), 103. https://doi.org/10.1038/s41597-019-0104-8

Pernet, C. R., Garrido, M., Gramfort, A., Maurits, N., Michel, C., Pang, E., Salmelin, R., Schoffelen, J. M., Valdes-Sosa, P. A., & Puce, A. (2018). *Best Practices in Data Analysis and Sharing in Neuroimaging using MEEG* [Preprint]. Open Science Framework. https://doi.org/10.31219/osf.io/a8dhx

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*(2), 127–152. https://doi.org/10.1111/1469-8986.3720127

Rodrigues, J., Weiß, M., Hewig, J., & Allen, J. J. B. (2021). EPOS: EEG Processing Open-Source Scripts. *Frontiers in Neuroscience*, *15*, 660449. https://doi.org/10.3389/fnins.2021.660449

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. *International Journal of Psychophysiology*, *156*, 18–39. https://doi.org/10.1016/j.ijpsycho.2020.06.016

Šoškić, A. (2021). Transparency in reporting on ERP research and how we can improve it. *Psychologie Und Gehirn 2021, 2-4 June 2021, Online [Programme Book]*, SP022. https://pug2021.de/ProgrammeBook.pdf

Šoškić, A., Gau, R., & Styles, S. J. (2022, September 12). *Fully transparent ERP & MRI study methodology descriptions with ARTEM-IS and eCOBIDAS*. INCF Assembly 2022, online.

Šoškić, A., Ković, V., Savić, A., Kappenman, E., Ke, H., & Styles, S. J. (2020, October 8). *Towards an Agreed Reporting Template for ERP Methodology: A draft proposal and roundtable discussion*. LiveMEEG, Virtual. https://doi.org/10.17605/OSF.IO/PDX6Y

Šoškić, A., Ković, V., & Styles, S. J. (2022a). ARTEM-IS: Improving Reproducibility and Replicability of EEG Research by Enhancing Transparency. *Acta Neurobiologiae Experimentalis*, *82*, 28. https://drive.google.com/file/d/1Da-jwzt1KZEAq1mjDhgmxtjhetp9fAKZ/view

Šoškić, A., Ković, V., & Styles, S. J. (2022b). Leveraging community expertise to improve methods reporting in EEG research with ARTEM-IS. *Psychophysiology*, *59*, S17. https://doi.org/10.1111/psyp.14139

Šoškić, A., Styles, S. J., Kappenman, E. S., & Kovic, V. (2022). *Garden of forking paths in ERP research – effects of varying pre-processing and analysis steps in an N400 experiment* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/8rjah

Styles, S. J., Kovic, V., Ke, H., & Šoškić, A. (2021). *Towards ARTEM-IS: An evidence-based Agreed Reporting Template for Electrophysiology Methods - International Standard*. PsyArXiv. https://doi.org/10.31234/osf.io/myn7t

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

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

Yang, Y.-F., & Styles, S. J. (2022, June 27). *Fully transparent ERP methodology descriptions with ARTEM-IS web app*. Society for the Improvement of Psychological Science, Vancouver. https://docs.google.com/document/d/1Rka3f0ZY9BBZVIlZoVINrpyaJYeitpq1yKXBLukeMng/edit