Guideline for Electroencephalography

1. Guideline for Electroencephalography (EEG)

1.1 Introduction

The document *COMFOCUS D6.1 Guideline for measuring psychophysiological responses*[1] contains detailed general guidelines and minimum reporting checklists for psychophysiology measures as well as specific guidelines and minimum reporting checklists for Heart Rate, Electro-dermal Activity, Eye-tracking & Facial Expression Recognition.

This current document builds on COMFOCUS D6.1 Guideline for measuring psychophysiological responses to present harmonized protocols for studies that use electroencephalogram (EEG) for implicit measures in food consumer science (FCS) context.

We recommend the researcher to first and foremost consult Chapters 2 "*General guideline for psychophysiological measures*" and 7.1 "*Minimum reporting checklist for all studies using psychophysiological measures*" in the COMFOCUS D6.1 Guideline for measuring psychophysiological responses. These chapters present an overview of common factors to be considered in the design and implementation of a study using psychophysiological measures Chapter 2) and good practice minimal reporting for publication/dissemination of studies using psychophysiological measures (Chapter 7.1).

Where appropriate, to prevent duplication, we refer to these documents in the manuscript below. We also refer to external published papers, documents or guidelines where appropriate.

This guideline discusses protocol- (section 1.2) and technology- (section 1.3) related factors that should be considered and reported in the study. We also provide recommendations on harmonised measures (section 1.4) and stimuli (section 1.5) for EEG studies. For convenience, we provide a minimum reporting checklist (section 1.6) that contains the factors that should be considered and reported in a study using EEG.

1.2 Protocol

This section describes main factors to consider when designing a protocol for a study involving EEG. For full information about the general study design factors and reporting standards please consult section 2.3 in *General guideline for psychophysiological measures*, for convenience, we have detailed what this document discusses below.

[1] Authors: Austėja Kažemekaitytė (UNITN), Philip Dean (UOS), Alejandra Bermúdez (IRTA), Wil van Dommelen (Noldus), Simone Cerroni (UNITN)

Good basic guides for EEG protocols are available from Farrans et al. (2019), Niso et al. (2022), Pernet et al. (2020) and Nebe et al. (2023). Farrans at el. (2020) in particular have a very detailed protocol guide we recommend adapting as a starting point for your own protocol.

We recommend reading these, along with resources such as:

- https://erpinfo.org/resources
    - See also event related potentials (ERP) CORE for Design & analysis: https://erpinfo.org/erp-core
- https://github.com/voytekresearch/Resources/wiki/Running-EEG
- The Turing Way has free guides for researchers guides such as their "guide for project design": https://the-turing-way.netlify.app/project-design/project-design.html
- Some technology manufacturer have some guidelines as well:
    - https://pressrelease.brainproducts.com/category/support-tips/
    - https://www.neuroelectrics.com/api/downloads/eeg_basics_principles.pdf
    - https://www.emotiv.com/eeg-guide/
    - https://www.gtec.at/plan-your-lab/
- Some Societies have lists of guidelines:
    - Society for Psychophysiological Research: https://sprweb.org/page/Guidelines_Papers
    - International Federation of Clinical Neurophysiology (IFCN) : https://www.ifcn.info/endorsed-guidelines.asp
    - American Clinical Neurophysiology Society: https://www.acns.org/practice/guidelines

1.2.1 Design

The following table (Table 1) presents EEG specific factors that should be considered when designing a study. Where applicable, information related to these factors should be reported in the final research output.

In the design stage it is of benefit to share project ideas with others in your research group/team or trusted colleagues within the research area or allied areas (neuroimaging is an area which spans psychology, biology, engineering, computer science, physics, mathematics amongst others). This not only ensures a good design from the outset, but also reduces the possibility of any bias, or limitation in your own thinking (Niso et al., 2022). A "premortem meeting" can also assist in design, where fellow researchers meet at design stage, assume the planned study has failed and work backwards to determine possible causes and control for them where possible

(Niso et al 2022). Sharing at the design stage is most useful, but sharing at any stage can help improve a project.

Table 1. Study design factors for studies using EEG

Factor	Recommendations
General Information **Pre-Registration** **Study design** **Synchronization** **Sample Size** **Documentation**	These headings/topics are covered in the section 2.3 in *General guideline for psychophysiological measures,* COMFOCUS D6.1 Guideline for measuring psychophysiological responses. Clear, detailed information on study, study design, stimuli and all forms administered must be given, and ideally pre-registered. Data sources and tasks must be synchronised / time-stamped in a way that allows analysis. ***Specific note for EEG*** The need for open science in EEG is discussed in Clayton et al (2022) and Garret-Ruffin et al (2021). A fully defined design, with preregistration and documentation of methods and measures used are particularly useful in EEG research, where there are a multitude of setups, preprocessing steps, analytical steps, measures/variables all involving numerous parameters. And all of these aspects interact and affect each other (e.g. task parameters used will impact what EEG measures you can analyse), so they ideally should be considered at the same time in project design. This is the "garden of forking paths" (as discussed in Niso et al. (2022) [from Gelman & Loken (2013)]). Where detailed design may be difficult (e.g. inexperienced researchers), Niso et al. (2022) suggest to ask for support from those with experience, and use a pilot study to define the finished design. A good starting point to look at ERP studies is ERP CORE: https://erpinfo.org/erp-core, which is a compendium of open

	resources & experiments, stimulus presentation scripts & analysis pipelines. A good EEG example of replicability is the #EEGManyLabs project (Pavlov et al., 2021).
Study setting	EEG can be acquired in static, wired setups in the laboratory or virtual (VR/AR) environments, or with mobile EEG setups within the laboratory, virtual environments or within more naturalistic settings (e.g., home, restaurant, office). Choice of setting will influence the ability to control for confound factors, the type and number of stimuli you present, the technology used, and the data acquired. The appropriateness of the technique for your research question, its interpretability, and the ability for it to achieve enough statistical power should be considered when choosing a setting.
Study environment	It is recommended that pilot studies are conducted to check whether the study setting creates any issues of data quality or data confound. The best way to reduce artifacts and noise in the data is at acquisition, as any post-processing to remove artifacts will always be partial, and can create artifacts of their own (e.g. filtering, Niso et al. (2022)). Farrens et al. (2020) developed a protocol with precise description of how to increase your signal to noise ratio (SNR, i.e. produce the cleanest data) and openly published it on protocol exchange. We recommend reading it, and publishing your own protocols in a similar fashion. Environmental conditions can affect EEG signals (temperature, humidity, rain), as can unrelated perceptual stimuli (e.g. audio noise). For this reason, lab experiments are usually done in a comfortable temperature, sound attenuated/quiet room,

	sometimes electrically shielded (see next section on electrical confounds). • Warm conditions cause participants to sweat, this saline water gets under the sensors (electrodes) and causes low frequency drift. • Warm conditions can increase feelings of sleepiness/fatigue, which may show in task performance and EEG signal. • Warm, low humidity (dry) conditions can dry out water and saline based EEG sensors quicker. • Cold conditions can cause muscle tension (see muscle noise issue) and discomfort. • Rain (when outside) or high humidity will affect EEG signal. • Noisy, or variably/sudden noisy environments will cause alterations in EEG signal. Non-neural noise: as EEG amplifies electrical changes on the scalp it can also amplify the much greater non-neural "noise" (e.g., electrical, magnetic, radiofrequency, movement, muscle, eye movement and blinking). This issue becomes greater if sensors have high impedance, where they can act like an antenna to the electrical and radiofrequency noise. Electrical, magnetic, electromagnetic and radiofrequency (WiFi, mobile signal) signals in the environment impact on EEG signal both in the data recorded and in interference/drop out of data transfer in mobile EEG setups. • Large electrical/magnetic noise artefacts are difficult to remove once part of the data. • Laboratory experiments try to reduce these sources of noise. For example, recording in an electrically and sound shielded room with DC (not AC) lights and as

	many electrical items (e.g., mobile phones, fans, electrical trunking or wiring) as possible outside this room. See Farrans et al. (2019) for a setup in this mould. ● Some newer EEG amplifiers with high input impedance, along with “active” sensors and shielded leads can be more robust to electrical interference. However, this can still be an issue and should always be considered in design. ● Mobile EEG setups can send the EEG data to the recording device via Wifi or bluetooth. ○ Radiofrequency signals in the environment can affect this transmission. Especially in areas with high usage where frequency bands used by e.g. mobile phones/laptops will change very regularly as users log on and off and so therefore start to use the same frequency band as your device. ○ It is worthwhile to check the radiofrequency environment you will be recording in, and pilot data to look for data dropout. If it is not possible to move to another location, it might be possible to use a frequency band outside that used in the WiFi environment. Movement of the participant (e.g. in mobile experiments) can cause changes in EEG signal, by either ● Moving sensor against or away from the scalp, causing large changes in signal due to change in impedance or scalp area. Movement of sensor leads and amplifier may also cause signal change. ● Resulting in muscle movement (forehead, neck, jaw, shoulders), which will cause high frequency artefact in the data. If you are interested in spectral frequencies in

	the same frequency as muscle noise, this may be problematic. Eye movements and blinking create large EEG signals compared to neural signals, and they can be largest at frontal sites where other signals (e.g. frontal asymmetry) are present. Long eye blinks or closing eyes for a short while can cause a burst of alpha frequency oscillations, especially when fatigued. Gamma band oscillatory activity is also affected by microsaccadic (small eye movement) spike artifacts in this frequency (Niso et al., 2022; Yuval-Greenberg et al., 2008). Monitoring eye movements and blinks enables this confound to be assessed. Eye movements and blinks that always/often occur at the same time as a stimulus or response will be very difficult to remove from analysis and make interpretation difficult. This might be the case when e.g. the stimulus of interest flickers on the screen or is a sudden change in luminance. Eye blinks also often occur when a participant makes an eye movement (at beginning or end of saccade), and blink rate is related to task difficulty (lower rate with more complex tasks). Therefore comparing conditions which have different difficulty or eye movement will mean differing eye blinks and eye blink correction. This may affect your data and should be considered in design and interpretation. Random (stochastic) eye blinks (timing unrelated to task) are easier to remove as most algorithms to remove eye blinks assume they are stochastic. Eye movements and blinking can be affected by: ● Changing light conditions / luminance of stimuli ● Dry, dusty or windy conditions (drying eye) ● Alertness ● Response - some participants (especially clinical populations) may look to their hands when responding. Complex responses (e.g. many options) will make this more likely. This is problematic as it always occurs at

	same time as response and so will be difficult to disentangle from neural response.
Study timing	Experiment timing: Fatigue/alertness can have an effect on task performance and EEG signal, both from an attention point of view (e.g. P300) and especially from a spectral/frequency analysis perspective, where greater alpha power can be seen in fatigued participants. This fatigue-related alpha is difficult to remove from the data once acquired. In this context, a subjective measure of fatigue/alertness would be recommended, and multi-timepoint studies should ideally be at similar times of day to account for alertness. Event timing: If the design has events (e.g. stimulus or response) or time-periods of interest then exact timing of these events/periods must be obtained (millisecond accuracy, synchronised to EEG signals). These "markers" or "triggers" should be labelled to allow identification of the specific event or period. Distinguishable labels for each event type/category of interest are needed to enable the analysis. When designing the study, think whether you need to know the start of the event, the start and the end, whether you need to know category type (e.g. Food vs Non Food) or sub-categories as well where they might impact (Food-Fruit, Food-Vegetable, or even Food-Apple, Food-Potato). If these are not recorded in the EEG data, it would be useful to record the exact stimuli shown in the behavioural task logfile so that any unforeseen issues can be found.
Study design **(Static vs mobile EEG)**	There are good mobile EEG setups currently available, but static EEG setups offer the most controlled setup and best data quality. Issues to consider for mobile EEG are: ● Head (& body) movement: any movement of the sensor against/away from the scalp will cause artifactual changes in EEG signal, as will movement of the sensor leads or

	amplifiers to some extent. This is also an issue in static setups, but is pronounced in mobile setups where the head and body is more likely to move. Whilst mobile caps are designed to reduce cable, amplifier, sensor and cap movement, some will still be present. It is important to ensure that the cap is secure, and sensors will not move away from the scalp. It is recommended to have a dynamometer in your mobile EEG amplifier/recording to take into account the movement of the head. ● Muscle artifacts: more movement of head, neck, shoulder, jaw, forehead will cause more muscle signal high frequency artifacts in the EEG signal. In lab studies, participants are asked to keep as still and relaxed as possible, and shown what happens when they e.g. clench jaw or raise eyebrows. For mobile EEG this is more tricky, but should be considered in design. ● Data drop: It is possible in a wireless setup that there will be missing packets of EEG data due to loss of communication signal (radiofrequency interference, distance to recording device, battery, computer processing buffer failure). Testing this with pilot data in an experimental setting is recommended. ● Sending event triggers: can be more problematic in a mobile EEG setup where the equipment is not physically wired together (stimulus PC - Amplifier - Recording PC). This is not an intractable problem, but may require planning. It may not be possible with some mobile EEG setups, where alternative forms of synchronisation may be required (see *General guideline for psychophysiological measures,* COMFOCUS D6.1 Guideline for measuring psychophysiological responses). ● Number and type of sensors: Discussed in the next row of this table as well. Mobile setups tend to have fewer sensors due to technical considerations of data transfer

	ability and cap setup. They also are more likely to have sensors such as dry electrodes (for ease/speed of setup), which have lower data quality. Ensure this is adequate for your experimental aims. ● Sampling rate: mobile EEG setups also tend to have lower time resolution capacity, due to technical constraints of data transfer, so number of sensors and resolution must be designed/piloted to ensure that IT buffer limit is not reached causing data recording drop-out/loss.
Study Design **(Sensor type)**	Sensors can be dry electrodes, water-based electrodes (e.g., MOBITA BIOPAC), saline nets (e.g. Brain Products R-Net) and gel-based electrodes. They can be passive sensors (just convey data to amplifier to amplify), or active sensors (have some sensor-level amplification). ● Active sensors are more expensive, but by amplifying the signal at scalp level they can reduce the influence of external noise and reduce setup time for gel-based EEG. Sensors can have leads that are shielded or unshielded to electrical noise. In general, dry electrodes still have substantial data quality issue trade-offs, and water based electrodes can also have impedance issues, especially as they dry. Saline nets can also have an issue with drying, but can be “topped up” during a task. They are typically used where setup needs to be quick (e.g. babies/infants) or where number of sensors is required to be large (e.g. 128 channels). Gel-based electrodes have been the standard, and although require some setup time (less if using active electrodes), they have the benefit of being able to reduce impedances to a consistent low level, and have more stable signal over time.

	How the sensor is set up also influences data quality: how the skin is prepared before gel is used (alcohol, nu-prep), the type of gel used (e.g. High Chlorine Gel), the type and size of sensor (e.g. Ag/AgCl ring electrode), how the gel is applied and how/to what level of accuracy the impedance can be checked (e.g. traffic light system only, or real figures).
Study Design **(Sensor number)**	Choice of number of sensors determines what sort of analysis you can do with the data. ● Single or 2-sensor setups (e.g. BIOPAC EEG) can be used to estimate global EEG changes in spectral frequency, or record site specific changes. ● Small sensor numbers (8, 12, 20, 32) will increase some spatial specificity of the signal and allow some ERP data collection along with spectral frequency. All these setups would allow "sensor-space" analyses. ● Even greater sensor numbers (64, 128) can allow some "source-space" (where neural basis is inferred) or EEG connectivity analyses. Sensor number will also affect choice of reference (average reference not advised on small sensor numbers). It is important in your design to understand what the choice of sensor number allows you to infer from your data, and what it will not allow you to infer.
Study design (randomisation and counter-balancing)	The eye, movement, muscle, electrical (and other) artifacts described above will become more problematic if they always or often occur at the same time as the task/choice/response of interest in your task. For instance, if you do not control for eye movements, and participants always move their eyes when they see the stimuli (especially if in same direction), or blink, or always look to hands (especially same hand) you will have a consistent artifact across

	all trials of the task that is larger than your neural signal of interest. And it will be difficult to disentangle. Your signal will be under the artifact, and any correction algorithms will likely take out some or all of the underlying signal of interest. If this eye movement varies between conditions, or participants, then differences between conditions/participants might be affected by these eye movements and/or their correction. In the extreme case of perfect overlap in one condition, but not another, you will get no signal in one of the conditions purely due to this artifact. ● Artifact correction techniques work best if the artifact is stochastic (ie randomly timed). So if you cannot control for the eye, movement, electrical (and other) artifacts, then you need to ensure they are randomly timed and not task-related. Neural signals, and behaviour, will be influenced by preceding stimuli, and with time across task. For this reason, it is recommended to have stimuli categories correctly randomised in a way that suits your experimental question/design. For instance (non-exhaustive list): ● Stimuli are completely randomised ● Stimuli are semi-randomised ○ e.g. so that similar proportions of condition stimuli are presented within the first half/second half, or for each quarter to control for fatigue/learning effects ○ e.g. so that there is equal transition probability between each condition, so the influence of preceding stimulus on current stimulus is controlled. ● Stimuli are randomised once, but always presented in the same fixed order. For example, randomised presentation

	order for equal transition probability and always presented in the same order to each participant, this can help with e.g. multi-timepoint analyses. You can create more than one "fixed order" and give this (e.g. order A) to the same participant to try and control for any remaining order effects. Neural signals can be influenced by response modality, even prior to response. For instance, always responding with e.g. right hand can cause lateralisation effects and may affect any measures of asymmetry. Handedness is typically taken in demographics for this and task-based reasons (as handedness will also influence task accuracy/timing). See Metzen et al. (2022) for discussion of influence of handedness of symmetry, especially frontal alpha asymmetry. Counter-balancing the hand used/buttons used/reponse used across participants or tasks can help control for this confound. If it cannot be controlled, then it needs to be noted and interpretation done in light of the confound. Counter-balancing any other potential confounds can also aid in reducing their effects on the EEG signal and behavioural response.
Study design (analysis)	The study/task should be designed with the aim/goal of the study in mind, and the analysis to be done. Exact analysis methods should be chosen while designing the task (Niso et al 2022), and ideally preregistered to prevent introducing biases such as p-hacking. For example, if you wish to analyse slow oscillations, then the time windows for analysis must be long enough to allow this and the data collected with appropriate sample rate and analog filter. See Technical Considerations section for more detail.
Study Design (hyperscanning)	Hyperscanning refers to obtaining simultaneous recordings from more than one participant and is useful for interactive and social studies. It has some technical and analytical specificities that need

	to be considered. For more information, see Barraza et al. (2019) and Brain products blog by Alex Kreilinger: https://pressrelease.brainproducts.com/hyperscanning-brainamp/
Study Design **(Secondary & simulated data)**	Increasingly, open EEG repositories are being used to disseminate research, and in study design it is worthwhile investigating whether you can use this data instead of collecting your own, to pilot your idea or analyses, to replicate, or to add as an extra condition in your study (e.g. control group). Simulating data might also help pilot your analyses and the effect of different parameters. They can be used for prospective power analyses. For example, to investigate the effect of cleaning electrical line noise in your data, or of a particular filter, on your measure of interest. There are resources to help you simulate data described in Niso et al. (2022).. Using secondary or simulated data also aids with sustainability and the aims of "slow science", as generating more data which is not needed uses up resources, time, and data storage. A discussion of sustainability in EEG is in Niso et al. (2022) (4.4), and discussion on sustainability in neuroscience can be seen in Rae et al. (2022). Where possible, planning to make your data open access is recommended.
Sample size	Sufficient sample size should be determined using power calculations. Published EEG studies can be underpowered, therefore researchers should be precise in setting the effect size and power level prior to calculating the sample size. As stated above, simulations/simulated data can be used for power analyses, or resources such as #EEGManyLabs or ERP CORE (Niso et al., 2022, see 3.4.2 in particular).

Measures	Feature extraction: Various features can be extracted from the EEG data, such as Event-related potentials (ERP), power spectra, coherence, and phase synchrony. These features can provide information about the frequency content and connectivity of the signal. Choice of measure will impact how you design your study, and the analysis to be done. It is sometimes necessary to choose between e.g. ERP analysis and spectral/connectivity analysis as a study design that incorporates both successfully may not be possible. In general, choosing a limited number of features to extract that are closely related to your hypothesis and have evidence for their use is recommended. More information on measures used in EEG can be found in the Harmonised measure section (1.4, Table 5)
Interpretation	There are prevalent logical fallacies within EEG research (See Niso et al. (2022), but also Sinnott-Armstrong and Simmons (2021)). It is important for a researcher to be aware of these and take steps to prevent them, such as (more detail in Niso et al., l 2022): ● List all theories or hypotheses that might explain a result. Overlooking one might lead to logical fallacy such as false dichotomy. ● Think about which theories or hypotheses are logically compatible with others. Could more than one be simultaneously true (or not distinguishable by your experiment)? ● Including stimuli or tasks for all variations within the category you are interested in. Not including a subclass of the category will lead to a fallacy such as hasty generalisation.

	When interpreting EEG data be aware of what the measure you are using represents (e.g. the ERP, frequency band), what assumptions are inherent to its use (e.g. is there a known biological basis/source, has it been linked definitively to a specific psychological construct), and what non-task factors can cause signal change (e.g. movement, impedance, pre-processing such as filtering can alter signal, different number of trials in an average, or confound factors such as fatigue). Interpret the data in light of all these caveats. In particular, EEG analyses can be prone to ascribe definitive psychological functions to measures that are not present, are based on assumptions, are not as clear (e.g. could be attention, memory, novelty), or are actively debated. And as discussed in the Sinnott-Armstrong and Simmons (2021) paper, some well known measures are based on research that is not replicable or has made some logical errors in its interpretation. A good design, with clear theory and hypothesis should enable an informative outcome despite these caveats no matter the experimental outcome. When analysing and visualising the data, bear in mind biases included in this (see section 1.3.1, table 4 for more details on how to avoid these).
Choosing appropriate comparisons	Choosing appropriate comparisons when analysing EEG data depends on the research question and the hypotheses being tested. For example, depending on setup, you might think to have: ● Control conditions: When comparing brain activity between groups or conditions, it can be important to include appropriate control conditions. Control conditions should be designed to isolate the effect of the

	manipulation of interest and control for confounding factors. • Baseline activity: It can be important to compare brain activity to a baseline measure. Baseline activity can be obtained by measuring brain activity during a neutral or resting state, or by using a pre-stimulus period as a baseline. Baseline correction in EEG analysis requires some "neutral" baseline. • Hypothesis-driven analysis: It is important to have clear hypotheses when making comparisons, rather than conducting exploratory analyses without a specific question in mind. This can reduce the risk of false positives and increase the interpretability of the results. • Replication: It is important to replicate findings across independent samples to ensure the reliability of the results.

1.2.2 Sample selection

- *Participant factors that influence data*

Table 2. Participant factors that influence EEG measures

Factor	Details
Demographics	Age, Gender, Race & Ethnicity discussed in *General guideline for psychophysiological measures,* COMFOCUS D6.1 Guideline for measuring psychophysiological responses (link)
Handedness	In addition to above demographics, it is recommended to record handedness of participants to control for effect on behavioural and EEG data. One measure that has been widely used (but also is not perfect) is the Edinburgh Handedness Inventory (EHI, Oldfield 1971).

Other Demographics	It is recommended to record these, but not exclude participants unless deemed necessary for your study. ● Eyesight ○ Visual acuity ○ Colour blindness ● Hearing (if audio task) ● Head size (recorded on day - circumference at widest part of head just above the ears, across the forehead, to the inion).
Clinical factors	It is recommended to record these, but not exclude participants unless deemed necessary for your study. Recording these has some ethical considerations, and good data management of health data must be evidenced. ● Epilepsy ● Sleep Disorders ● Medications ● Ongoing treatment ● Mobility (for mobile EEG) ● Arthritis (for response tasks) ● Heart disease (for mobile EEG or active tasks) ● Any other issue
Diversity	EEG in particular has been predominantly performed on WEIRD (Western, Educated, Industrialised, Rich, and Democratic) populations, and this needs to be factored into both previous interpretation and future experimental design/recruitment (Heinrich et al., 2010). This lack of diversity is also seen in researchers at institutes and those that get funding and time to do research (Niso et al., 2022).

Differences in hair/scalp and skull shape.	There is an issue with racial and phenotypic bias (Webb et al., 2022, Choy et al., 2022, Parker et al., 2022), where devices and techniques themselves are built for WEIRD groups without insight into its effects on data from those in other groups (e.g. skin colour, hair type), and possible exclusion of those participants. • Some groups have coarse, curly or thick hair (or braids, twists, weave, hair extensions, wigs) which can have large effects on data quality (impedance) in conventional EEG setups, and so are often excluded either at screening or in analysis. • Hair type/style might also mean participants will need to plan taking part in experiments with e.g. gel onto scalp (and then washing hair). GIving participants information on what exactly the procedure will involve will allow them to plan and take part if wanted later. • EEG with caps to keep electrodes in 10-20 positions often are modelled on European head shapes, and are not as well fitted to e.g. Asian head shapes (although there are companies that can supply different caps if needed). • Source modelling of EEG signal is sometimes based on brain structure models from WEIRD groups, and so not representative across cultures. Neuroimaging (MRI) has taken steps to get better representative average brains for ethnicity and age groups which can be used. There are steps within the research community to create technology to help (Etienne et al 2020, https://hellobrainlab.com/research/eeg-hair-project/) and guidelines on how to record data in this context (e..g https://pressrelease.brainproducts.com/inclusivity-in-eeg-research/). It is recommended to read the guidelines above, recruit as diverse a population as possible to allow generalisation of results, adapt

	techniques to allow good data quality from all participants, and where participants have been excluded for reasons above, to note this in the report/publication
Behavioural factors	Ask participants if they have experienced any recent change from typical (for them) level of sleep, food schedule, caffeine (chocolate, cocoa, coffee, tea, caffeinated drinks, caffeine pills) & alcohol. Take note of large alterations in these (e.g. lack of sleep) recently, or ask to come back dependent on effect on study Before setup, ensure that the participant does not need to use the bathroom, as procedure/setup can be long.

- *Participant factors that influence safety*

Generally, EEG is a safe method, with discomfort and sensitive skin being some of the risk factors that should be monitored (see inclusion/exclusion section below). Discomfort with EEG cap can be due to head size/shape and size of cap. Ideally the cap should be tight to get a good contact and prevent moving, but not too small for the participant (causing discomfort and meaning electrodes in incorrect areas). So head size and cap size are recorded.

Certain participant characteristics can impact on the data (see table above), and clinical characteristics such as epilepsy, seizures and some sleep disorders (such as sleep apnoea) may be seen in the data. EEG has also been used in medical settings to look at brain tumours, brain injury/concussion, dementia, and encephalitis (brain inflammation). Medication use and medical treatment has the potential to influence EEG signal, but holds no safety issue. As discussed in ethical factors below, unless the research institute has a specific medical setup, then diagnosis is not typically possible as researchers are not qualified and equipment states "not for medical use".

There are factors related to the technology that need to be considered for safety. All EEG amplifiers should have CE marking (or equivalent) for safety. Any use of external equipment needs to be done with care and discussion with technology manufacturers to ensure safety. Unless stated as safe in the manual/by manufacturers, the EEG amplifier should be disconnected from the mains (on battery power) before cap or sensors are put on the participant. Neuroelectrics setup (and other setups - check your manual) should not be switched on/off when on the participants head, and not set up when connected to the power network.

- *Inclusion/exclusion criteria*

Participants with sensitive skin may experience discomfort with EEG gel, or alcohol. In addition, use of water or saline water electrodes might cause dripping into eyes and discomfort of sensitive skin. Any issue of minor rash, cuts or scratch would also potentially cause discomfort. Recommendation is to inform participants in the information sheet of these factors, then include them if they consent to take part.

Example wording:

"Those with sensitive skin may experience discomfort when the gel is applied to the scalp and face, please notify us if you have especially sensitive skin. There is facial moisturiser to use after the experiment. The gel is hypoallergenic. However, if redness at site of gel application has not gone after a week, consider contacting your GP."

Any participants with more prominent rash (e.g. eczema), open wound or sore on head at sites of electrodes or cap should be excluded until these have healed due to comfort and hygiene issues. These can be asked at recruitment and also checked and noted on arrival of participant/setup of cap.

1.2.3 Ethical factors

For general psychophysiology ethical factors, see *General guideline for psychophysiological measures,* COMFOCUS D6.1 Guideline for measuring psychophysiological responses.

Psychophysiology measures may fall into categories of *sensitive* or *health* data which require ensuring certain levels of data protection, and it is possible to collect *identifiable* or *special category* data, so a data management plan must be created. Researchers must be trained in the techniques to ensure efficient, hygienic and safe setup resulting in good quality data. Participants must be made aware of the ability to withdraw at any point and be explained the technique both before and during setup (before each step).

First, an interesting ethical point raised by recent open science and sustainability initiatives. Open EEG (and other modality) repositories already exist with a number of projects across substantial numbers of participants in a number of settings (see Niso et al., 2022). Resources also now exist to simulate EEG data using intricate models (Niso et al., 2022). As these datasets and simulation models exist, a researcher must first ask themselves: do I need to collect my own

data, or can I use existing data or simulations? Either totally, partially (e.g. as a control group) or as a pilot study/proof of concept. To do so makes use of already existing data (sustainable use of resources), and prevents collecting more data where it is not needed (wasting participant time and research money). In addition, not making your data open will also reduce its sustainability. See also Rae at al. (2022) for further comment on sustainability.

Whilst EEG has the potential to be used as a diagnostic tool, typically EEG amplifiers are intended for research purposes. They are not medical devices, and use for diagnosis is strictly forbidden. In addition, experimenters are typically not trained in interpreting EEG data in a clinical way, and as such, EEG used in this lab setting cannot be used as a diagnostic tool. However, this would need to be made apparent to participants, and if the research institute had appropriate medically trained staff, then it could be sensible for this to be checked for each participant and a procedure followed (e.g. take GP details in recruitment and contact GP, or ask participant to).

Information sheet: Within the information sheet, it is useful, depending on setup to add these items (only use those that apply):

- *"We ask you to not wear any hair products (gel/spray) on the day of the experiment, as this may interfere with the equipment."*
- What sensors are going to be placed where and how many (e.g. on face/near eyes, on scalp, on/behind ears, in cap, or even on back (for ECG)).
- Whether alcohol, gel, saline, water will be added, and whether they can wash their hair afterwards: *"You are then given the opportunity and facilities (washbasin, shampoo, towels and hair dryer) to clean the gel from your hair."*
- Some people experience mild discomfort whilst wearing the cap. In rare cases this may cause a headache.
- Within the experiment, we may ask you to blink and move your head and eyes only at specific times as these processes affect the recording of the electrical signals. This may become tiring, but sufficient breaks are given to rest your eyes.
- Sensitive skin: *"Those with sensitive skin may experience discomfort when the gel is applied to the scalp and face, please notify us if you have especially sensitive skin. There is facial moisturiser to use after the experiment. The gel is hypoallergenic. However, if redness at site of gel application has not gone after a week, consider contacting your GP."*
- As we are using gel/cap/saline on the scalp, we ask for participants not to come if they have any major cuts, rashes, sores or open wounds until these have healed as they may cause discomfort.

- Specifics about sensitive, health data, personal data, special category data, and the data management of this.

1.3 Technological factors

In this section we describe the main technological factors to consider when designing and running a study involving EEG.

Table 3. Technological factors that influence EEG

Factor	Details
Equipment factors **Set-up factors** **Data**	These headings/topics are covered in the section 2.3 in *General guideline for psychophysiological measures,* COMFOCUS D6.1 Guideline for measuring psychophysiological responses. Sufficient information must be given on the technology, software specific settings of technology and software, the setup (e.g. sensor, screen, environment, baseline, instructions given) used to enable direct replication. Measures to be used in analysis must all be stated, explained in enough detail to replicate, and ideally pre-registered.
Hardware	Amplifiers have an inherent analog filter (low pass/high pass) that needs to be reported and accounted for in how it affects the raw data and the ERP/frequency bands you are interested in. It may also influence the ability to assess low frequency artefacts and remove high frequency artefacts (by under sampling). There is some discussion on how to minimise differences in multisite studies with different hardware: https://pressrelease.brainproducts.com/multicenter-studies/
Sampling Rate / Frequency	Sampling frequency used affects both the ability to see certain ERP/frequency bands and the ability to remove artifacts. ERPs, even the Contingent Negative Variation (CNV) will have an underlying period/frequency that needs to be properly sampled, and can be filtered out by hardware (see above) or sampling rate. Sampling rate needs to be high enough to get a good sampling of the 0-360 phase of the wave. A 1Hz wave will take 1000ms (1s) for a full 0-360 period, a 40Hz wave will take 25ms (0.025s). Sampling rate needs

	to be set to *at least* 4 times that of the highest frequency to be analysed (so 4 data points across the full 0-360 wave). Sampling rate can also cause the inability to remove high frequency artefacts, as they are not sampled enough (undersampling) and cause aliasing in the data. For example, specific artefacts such as Magnetic Resonance Imaging (MRI) or Transcranial Magnetic Stimulation (TMS) pulses will not be cleanly seen at lower sampling frequencies so will have a large effect on the data without being able to remove them.
Choice of Reference	Choice of reference is important, as it can affect your signal. See general discussion (https://pressrelease.brainproducts.com/referencing/), as well as discussion of this for Frontal Asymmetry in Smith et al 2017. Choice depends on where you expect your signal to be, what sort of signal it is, and also how many sensors (electrodes) you are using. Small number of electrodes typically means that average reference (sometimes called Common Average Reference [CAR]) is not feasible. Where CAR is inherent to the amplifier or software in small electrode arrays (e.g. <64 sensors), it is recommended to re-reference the data to specific reference electrodes in analysis steps. If you will be pooling sensors into areas, remember that the more sensors you pool, the closer you are getting to the average reference (i.e., pooling 32 sensors in a 32-sensor setup with average reference will result in a flat line [the average]). So choose pool sizes carefully in average referenced data. Using physically linked earlobe or mastoid sensors during acquisition is not recommended, as they are not a neutral reference and can introduce distortions (Pernet et al., 2020). If they are physically linked, you cannot determine the signal quality or impedance of the sensors independently, so they cannot be disentangled and controlled for.
Stimulus Software	Even if EEG data is synchronised with stimulus software (or time-stamped), there will still be variability in the timing of stimulus

	presentation (e.g. code says it presents stimuli every 3s for 100ms, but in likelihood this will not be exactly right each time), response registration (again, might be delayed due to device used, connection, software), and event trigger timing (some variability in time sent to EEG recording device). This accuracy will depend on software, hardware and task, and needs to be investigated (e.g. software says variability is +/- 5ms, screen refresh rate is 60Hz with 5ms response time), and some software gives an idea of its accuracy/jitter in the logfiles. However, this should also be measured individually by researchers to ensure accuracy (see Niso et al., 2022). The main recommendation is to properly measure and document your setup, and any differences in timing/internal clock/sampling rate that cannot be corrected between modalities (e.g. task, EEG, eye tracking, audio, visual, see Niso et al., 2022). See also Brain Products discussion: https://pressrelease.brainproducts.com/timing-verification/ This calibration of your setup is valuable and can represent research data in its own right to document, detail and ideally make open source to help others.
Event Timing	As discussed in the Design section (1.2.1, Table 1), and the Stimulus Software row above, timing of task events, and synchronised timing between modalities (if using more than one measurement) is crucial. Direct hardware connection can achieve this with appropriate software. Alternatively, introducing external triggers that can be recorded across all modalities may be a solution, repeated at intervals should internal clocks of each modality drift. A free, open-source software solution to synchronise multiple modalities is Lab Streaming Layer (LSL). See also Brain Products discussion: https://pressrelease.brainproducts.com/timing-verification/

Monitor	The type and refresh rate of the monitor can influence the timing of your stimuli, and in a modality that is accurate to the ms, this can affect data. A 60Hz screen refresh rate means it refreshes every 16.7ms. a new "image" cannot be written by the software until the screen refreshes. So if your stimuli is meant to happen at 100ms from start of trial, and the screen has just refreshed at 99ms from start, then it will actually be presented at 115.7ms from start. Whether this matters depends on the task, but it should be considered. There is also "response time" of a monitor, which means the time it takes to react to a specific input. If this is 5ms, then it will display the stimulus 5ms after it is asked to. Higher refresh rate screens can be obtained: use of gaming monitors with good graphics cards and response times can mean stimuli timing is not as much of an issue. Flat screen monitors are usually used, replacing older CRT (Cathode Ray Tube) monitors where pixels were "written" one at a time from top left to bottom right.
Recording Software defaults	EEG Recording softwares often have default settings, and their effect on the data needs to be understood. In particular, the sampling rate, resolution, range, the use of any filter on the data (high/low) at a group or individual level, the incorrect or inconsistent assignment of series resistance at sensor level, or sensor label at channel level will affect the data recorded. In addition, it is recommended to always record raw data. When setting up a recording workspace, you can set: ● What amplifier to connect to (this may affect what the software does to the data, and available options) ● Where the data is recorded ○ Any name prefix for data ● Number of sensors used ○ Sampling rate

	○ Resolution (in µV) ○ Low cutoff frequency (e.g. 10s or 0.1Hz, or "DC") ○ High cutoff frequency (e.g. 1000Hz) ○ The above then determines the range of the amplifier (+ to - mV) ● Any Ground or Reference series resistor (active sensor) ● The setup of the sensors for each channel ○ Names: a list from 1 to n of which sensor relates to which recording channel (e.g. channel 1: Ground, Channel 2: Fp1,..... Channel n: Reference). ○ What type of sensor each of these are (e.g. EEG, BIPolar, AUXiliary) ○ What physical channel these are ○ Any sensor-specific resolution, low, high cutoff, series resistance (active sensor). Usually all channels/sensors would be the same, but occasionally you may need to change this e.g. in case of ECG channel included which has resistor. ● Electrode Position File: where each electrode/sensor is in space. Typically, international 10-20 positioning is used. There are usually default files that come with the amplifier/cap, and it is good to check how they are configured and how they compare to standard montages. ● Filter settings: enable low and high cutoff filters, as well as notch filter (for electrical line noise at 50 or 60 Hz) for all sensors, or just some: ○ Raw data saving filters ○ Segmentation filters ■ Used when segmentation specified ○ Display filters ■ To clean data when monitoring during experiment, but raw data recorded without filter.

	○ **It is highly recommended that you record data without filtering, due to the potential effects of filtering on your data, and that it allows more transparent sharing of open data. So "raw data saving filters" are not recommended.** ● Segmentation or Averaging ○ Segment and average your data based on markers / event triggers. Can be saved in parallel with raw data ○ Can be used to see if a visible ERP or MEP (motor evoked potential) is formed with/without saving the data as a task/quality check. ○ **It is highly recommended NOT to *only* save the segmented or average data. Always save the raw data to enable sharing of data and also data quality check.** ● Some software allows different file formats to be saved, but a lot of software has its own bespoke format. ○ It is recommended for open data to record to standardised file formats, such as the European Data Format (EDF). Alternatively, to convert to this format after recording for data storage & analysis.
Recording Hardware	The laptop or PC used to record the EEG needs to be considered, as the signal recorded from EEG can be high resolution (1000Hz or every ms), across many channels (e.g. up to 128), with synchronisation with task (event triggers) and perhaps simultaneous recording of other modalities (AUXiliary, or Psychophysiology). This may require a lot of processing power to ensure data is recorded without any data drop. Recording software usually give specifications needed, as well as other setup requirements. For instance Brain Vision Recorder asks that you "make sure that the performance of your system is not impaired by background processes or other real-time applications which run with higher priority." by:

	• having minimal other software on the PC/laptop, only what is required for EEG recording. • running other critical near-time application on another PC/laptop (e.g. NIRS, eye-tracker, stimulation, voice, video, audio recording, flash application, time tracker) • stopping background programs/services when recording (system, indexing, scanning (anti-virus), updating (web browser, office, antivirus, one-drive). Do not browse the internet, or play multimedia files. • changing Windows so that screen saver, defragmentation, windows update and power management will not run/interfere with recording. • Recording to hard drive, not to online server (e.g. One-Drive) for updating reasons mentioned above. • (Can disconnect from WiFi/ethernet to ensure no update) The reason for these precautions is that recording runs with a "buffer" (which you can see during recording), and when a background process runs, taking resources, then this buffer is reduced until the program is no longer able to record data as it comes in, and data loss occurs. It may not be obvious that this has happened at the time, or in the data, but it will affect your data quality. Note that mobile EEG setups are sending this data typically across a Wifi or bluetooth connection, so may have more issues in data loss (from connection loss and processing limit reached) and buffer over-run dependent on setup. It is recommended that researchers use a setup as close to the above (minimal background processes) as possible, and trial their setup to ensure that there is no buffer over-run and data loss.
DC Recording	A note on DC (Direct Current) EEG recording, which can be specified in setup as the low cutoff filter ("DC") in certain amplifier and software setups. Non-DC amplifiers typically have a low-cutoff of some degree (e.g. 0.1Hz as above), meaning there is some spectral data loss below

	this, and some data alteration of the recorded data from the low-cutoff filter. DC setup allows recording of lower frequency spectral data (no low cutoff). The caveat is that when recording with DC, then the amplifier is more likely to hit the limit of its recording range (see above) due to the "drift" of the signal. Once it does this, the sensor (electrode) in question will flatline and no more data will be recorded. Electrodes with worse connection or impedance will have more drift, and so be more prone to flatline. Good setup will limit the likelihood of flatline. To counteract this, the researcher can "DC Correct" the data online during recording. Software with DC recording will tell you how close each electrode is getting to its limit of range, and you can press the "correct" button before this happens. This correction causes a large reset across all electrodes (so a lot of noise), and will have to be timed to be in a part of the data/task where this will not be an issue. Alternatively, you can set the correction to automatic, and it will automatically do it when a limit is near, but as it is automatic it can be any point in the task.
Impedance & Data Quality	As discussed in "study environment" Table 1, section 1.2.1, It is very important to acquire the cleanest (highest signal to noise) EEG data possible. Analysis can clean the data, but even filtering will have an effect on data. Impedance measured in setup refers to the resistance of the sensor-skin interface and can affect the quality of EEG recordings. Whilst scalp abrasion & conductive gel was used to reduce impedance traditionally to a threshold of <5kΩ, recent setups use high input impedance amplifiers to ease setup, and this can be used with e.g. active sensor, saline net sensors, or dry sensors. However, these high input impedance amplifiers can have increased skin potentials and other low-frequency noise/drift, especially in a warm environment.

	Impedance of the ground and reference sensors ideally should be kept low for good recording. Mismatch between reference and active sensor can cause increase in non-EEG signals in data. As discussed above in the design section, factors such as muscle tension or movement, sweating and nearby electronic devices can affect the data. Farrens et al (2020) developed a protocol with precise description of how to increase your signal to noise ratio. It is recommended when setting up EEG caps or sensors to go over all sensors once and ensure they get some signal (e.g. you clean with alcohol, abrade skin and add gel), then go back over them to improve signal quality and impedance. Once all sensors are initially prepared, it is easier to lower impedance. It can sometimes take a little time for the signal to "stabilise" in gel or saline caps at least, so after setup, monitoring the raw signal to check it is clean and consistent is also recommended before starting recording
Source localisation & Sensor placement	For some setups, it is useful to have the exact location of the sensors on each participant's head relative to physical markers on the head (nasion, inion, jawbone). When combined with an average brain structure or participant's own brain structure then a better model for source localisation can be made. There are many ways to achieve this, for instance with a Polhemus stylus to collect a 3D map of electrodes.
Calibration	EEG does not need to be calibrated each time in the same way as eye-tracking. However, some steps on initial pilot or on each recording will aid data quality. When piloting, checking that the event markers come through, and that the timing is correct/synchronised is important. In addition, recording data and analysing it to check for any systematic noise or artifacts (electrical noise, data drop) would also be recommended. Highlighting and combatting any unforeseen issues is possible at this early stage.

	On each setup, it is recommended to check that the electrodes are wired in correctly to the right channels (if this is an option in your setup), that the impedance is sufficiently low across electrodes, that the ground and references have good impedance, that the raw signal looks clean, free from drift and noise. Showing the participant the raw data and how it is affected by movement (frown, clench jaw, move head, move shoulders) and eye movement, closing eyes and eye blinks can be used to help get them relaxed for the task.

1.3.1 Data processing

Table 4 provides a list of data processing factors to be considered and reported for studies using EEG.

Table 4. Data processing factors for studies using EEG

Factor	Details
Analysis **Data transformation** **Data quality**	These headings/topics are covered in the section 2.3 in *General guideline for psychophysiological measures,* COMFOCUS D6.1 Guideline for measuring psychophysiological responses. Report how analysis variables were calculated from raw data, any transformations, the number of variables used. Monitor data quality thoroughly during acquisition and in analysis. Reporting measures of data quality is advised.
Data Quality	See section 3.4.3 in Niso et al. (2022) on Data Quality For data harmonisation, quality control measures, such as visual inspection of the data and consistent artifact detection, should be applied across different devices and manufacturers to ensure the validity and reliability of the data. See also: https://pressrelease.brainproducts.com/multicenter-studies/ In general, studies should report information on the data quality (signal to noise ratio, impedance, "raw data inspection" for flatline or extreme

	values) and pre-processing steps (number of artifacts found & removed/corrected), blink & eye movement parameters (e.g. in ICA (Independent Component Analysis)).
Analysis Reproducibility	EEG (and other modality) analyses are now so complex that it is difficult to detail exactly in methods sections, and most robust analysis is done through coding within the software (Niso et al., 2022). This leaves the code as the ultimate documentation of the analysis done and it should ideally be made open to the reviewers and the readers of a publication. For this to be useful, the code must be clearly written, well organised, documented and annotated for others to follow (with e.g. unused code deleted). A bonus is that this will also make it easier for the researcher to re-use or adapt. Use of version control systems such as git help with this. With some training and work, a pipeline can be made in code from analysis to "notebook" to draft paper using systems such as R Markdown (also possible in MATLAB/Octave and Python).
Analysis Software	See Niso et al. (2022) section 3.2.2 and Box 3 for discussion on typically used software, its merits and its caveats. An important note is that some EEG analysis software are more suited to open-science, in that they are easier to share code openly for reproducibility (see point in row above), users can see source code and create/adapt toolboxes, and some are freeware as well. Of the ones discussed, EEGLAB and FieldTrip are widely used and code can be shared, they run in MATLAB (which is not freeware, but Octave is [check compatibility]). MATLAB is also the environment that automatic analysis (*aa*, see row below) is built in. MNE-Python has the above positives, but is newer, so not yet as widely used, and runs in freeware (python). There are also other licensed softwares, such as Brain Vision Analyzer or BESA Research (ARTEM-IS (Styles et al 2021): https://artemis.incf.org/ has a further list).
Processing steps	Preprocessing can involve steps such as re-referencing, down-sampling, raw data inspection, filtering, blink correction (e.g. using ICA),

segmentation, artifact rejection, fourier transformation (spectral analysis), averaging, baseline correction. After this, things like coherence, connectivity, power, phase, amplitude and latency analysis can be performed.

Understanding how each processing step affects the data, and the changes they can bring is essential to EEG analysis. For instance, filtering can cause distortion of the EEG signal, shifting latencies, and even artificial peaks and oscillations. Understanding how and why can help determine the steps and parameters you use.

Checking the robustness of your analysis by looking at different options within a processing step can also help troubleshoot the utility of the step. If you do this, this needs to be reported in the publication for transparency. See also *multiverse analysis* in Niso et al. (2022), where a number of different "paths" in the garden of forking paths of analysis are applied to the same data to see their effect. Some analysis has already been done in this area, and may give insight and inform your analyses.

The order or connection of different processing steps can also alter data, and researchers need to be aware. For instance, time-frequency maps should be calculated before averaging across trials for induced responses (Niso et al., 2022). Prior pre-processing steps may preclude subsequent ones.

For harmonised data across sites, a standard pipeline of pre-processing and analysis is recommended.

The use of BIDS data structure (see section 1.6) enables the use of some recently developed attempts at standard pre-processing pipelines:

- Automatic Analysis (*aa*) https://automaticanalysis.github.io/ & https://github.com/automaticanalysis/automaticanalysis/wiki. This also allows full pipeline analysis (preprocessing - feature extraction (e.g. source-reconstruction, ERP, power & connectivity analysis and statistics)) to be coded. It is written in MATLAB and supports parallel processing.

	• PREP: http://vislab.github.io/EEG-Clean-Tools/ • HAPPE: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5835235/ • BEAPP: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6090769/ • AUTOMAGIC: https://www.biorxiv.org/content/10.1101/460469v2 • CTAP: https://peerj.com/articles/cs-108.pdf • Makoto: https://sccn.ucsd.edu/wiki/Makoto's_preprocessing_pipeline These are being updated and changed rapidly as the field progresses, so we cannot recommend any individual software. However, aa is the most complete, versatile and updated at the moment.
Filtering	Filtering is often used in preprocessing to remove unwanted noise and clean the signal of interest. You can also use filtering to extract a frequency of interest (e.g. band-pass filter to extract alpha frequencies). As discussed before, filtering has effects on the remaining data as well. Filters also attenuate signal around the cutoff (e.g. a 30Hz filter will also reduce signal in 29, 28, 27 Hz etc depending on the filter slope setup). Typically, you would filter out the unwanted high frequency data as muscle noise (e.g. 30, 40 or 50 Hz high cutoff). This might not be possible if you are interested in frequencies at this range, or those at the cusp of the filter. You would also use a low cutoff to filter out drift caused by e.g. perspiration, e.g. 0.1 or 0.5Hz. Some ERPs might be affected by low cutoff filters, and frequency analysis at this range would also not be possible. Lastly, if needed, a “notch” filter which cuts out the electrical line noise (50 or 60Hz dependent on country) can be used. Where the high cutoff filter is lower than the notch, this is not necessary. One point to note, the literature uses both high cutoff/low pass and low cutoff/high pass to describe these filters interchangeably. We have

	grouped the names that mean the same thing together above. Low cutoff filters are also sometimes stated in seconds in some software (e.g. 10s = 1/10 or 0.1Hz).
Data Visualisation	Carefully look at your data as much as possible during analysis. From raw data, and at each stage of processing to see what has happened and ensure it is what you need. This can be achieved through code plotting these stages for you to look through. Time-frequency analysis to look at spectral power, and topographic plotting to explore spatial patterns of activity across the scalp can also give insight into any issues.
Design, Preprocessing and Statistical power	Some preprocessing steps affect the power and scope of the analysis. For instance removing an electrode as it is too noisy or flatline will affect some analysis and affect harmonisation across participants/sites. Linear interpolation of the electrode will allow the analysis and harmonisation, but will affect the statistical analysis, as this data is based on other sensors and hence will reduce electrode variability and change patterns for the participant. A more common issue is artifact removal (removing e.g. muscle activity, eye movement, flatline), where a trial (or in some cases an electrode in a trial - this is not recommended) is removed as an artifact. This reduces the power of the analysis, with more trials removed as artifacts resulting in further reduced statistical power. Where this removal affects the balance of trials between conditions it can also lead to spurious results based on power and variability. This is because grand average EEG signals can differ purely based on the number of trials used. It is recommended to have more than enough trials in your task so that when artifact removal has taken place the data is still adequately powered. It is also recommended to report the trial number per condition used in the analysis, and balance the trial number in

	conditions as much as possible (it does not need to be exact, but large disparities in proportion are problematic).
Statistical Analysis	All analyses done on the data need to be reported (not just the relevant/significant ones) and ideally pre-registered. Also from Niso et al. (2022): ● Always account for multiple comparisons in mass univariate data analyses (like EEG). This can be something like permutation analyses, false discovery rate (FDR) or in simpler cases, Bonferonni correction. ● Take care when using powerful methods like cluster-based correction for multiple comparisons. These should not be used to localise effects in time or space. ● Avoid circularity or "double dipping" of data analyses (also see Pernet et al., 2020), for example selecting specific channels or source level regions of interest (ROI) based on grand average differences between conditions/groups and then performing statistical tests on these. These should be specified a priori based on previous literature or on independent data or statistical contrasts. ● To prevent increasing number of factors (e.g. in ANOVA), leading to increase in familywise Type 1 error rate, collapse/pool data across sensors (electrodes) and do not include all sensors as a factor, unless individual sensors are important for your scientific hypothesis.
Machine learning	This data-driven approach to analysis should have the same caveats and guidelines mentioned above for more hypothesis driven research. The researcher needs to be aware of the task, the recording and the assumptions and interpretations that are possible in this context. Avoid violations of strict separation between training and test data. See Niso et al. (2022) Section 3.5 for more guidelines.

Data Presentation	As discussed in Niso et al. (2022), Cooper et al. (2021), and Franconeri et al. (2021), the way data is presented influences researchers' inference and is affected by cognitive biases. These papers have recommendations on use of e.g. perceptually uniform colour scales over rainbow scales, and the importance of using box and density plots to show the underlying distribution of data.

1.4 Harmonized measures

There are some standard terminology and definitions in relation to data acquisition, which for reproducibility and clarity reasons we recommend researchers use in their metadata, but also in reporting/publishing data. Kane et al 2017 has a nice glossary of a number of commonly used terms which are of use to the researcher. However, for the below, we refer the reader to Pernet et al. (2020) Box 1 for more details:

- Session: Group of all neuroimaging/behavioural data from participant arriving until they leave
- Run: Uninterrupted period of continuous data acquisition without operator involvement
- Event: An isolated occurrence of a presented stimulus, or a participant response recorded during a task.
- Trial: Period of time includes a sequence of one or more events with prescribed order and timing. The basic repeating element of an experiment.
- Epoch: Outcome of data segmentation process e.g., segment time-locked to a particular event.
- Sensors: Physical objects/transducers that perform the recording (i.e. EEG electrodes)
- Channels: Digital signals that have been recorded by amplifiers
- Fiducials: Markers in well-defined location used to facilitate localisation/co-registration of sensors with other spatial data. Typically places at known location relative to or overlying anatomical landmarks.
- Anatomical Landmarks: Well-known, easily identifiable physical locations on the head (e.g., nasion & inion).
- Sensor Space: Representation of data at the level of the original sensors.
- Source Space: Data reconstructed at the level of inferred neural sources that presumably give rise to the measured signals.

In addition, there are some standard terminology and definitions in relation to data analysis, which for reproducibility and clarity reasons we recommend researchers use in their metadata,

but also in reporting/publishing data. We refer the reader to Pernet et al 2020 Box 2 for more details:

- Event-related Potential (ERP) component traditionally refers to a functional brain process that has a characteristic spatial distribution and canonical latency. Where this is not the case, deflection can be used as an alternative to component.
- Oscillation: term specific to a spectral peak within a frequency band of interest and not a general increase in MEEG power within a canonical frequency band. Defined by its peak frequency, bandwidth and power.

Table 5. Harmonized measures for studies using EEG

Factor	Details
Event-related Potentials (ERP)	Always label ERP by polarity and latency in ms in full (e.g. N170 or P300 [not P3]), this avoids confusion with older versions which were named by canonical order (P1, P2, P3). In addition, add the sensor/recording site to the name (e.g. vertex N100; FCz N100). ● This includes using P300a and P300b (instead of P3a, P3b or Late positive component [LPC]) ● Other terms refer to specific responses elicited in particular types of paradigm or to presumed mental states (for example CNV [contingent negative variation], MMN [mismatch negativity], ERN [event-related negativity]) Electrodes used and Latency window to quantify ERP should be explicitly mentioned.
Spectral / Frequency Analysis	Always report spectral frequency bands used. Canonical MEEG frequency bands: ● infra-slow: < 0.1 Hz ● delta: 0.1 to < 4 Hz; ● theta: 4 to < 8 Hz; ● alpha: 8 to < 13 Hz; ● beta: 13 to 30 Hz; ● gamma: > 30 to 80 Hz
Frontal Alpha Asymmetry	Difference in alpha power between the left and right hemispheres of the brain. See Smith et al (2017) for further details. See Metzen et al (2022) for discussion of influence of handedness on asymmetry and frontal alpha asymmetry.

Connectivity	This term covers many methods, which creates confusion. In general, it is “analysis that aims to detect coupling between two or more channels or sources” (Pernet et al 2020). Recommend to state whether connectivity done in sensor/channel or source space and explicitly refer to: ● Functional (correlational) connectivity ● Effective (causal) connectivity Pernet el al (2020) states that effective connectivity (causal) can only be obtained at the source level (not sensor or channel), after biophysical modelling, and considering volume conduction and spurious connections. Connectivity at the sensor/channel level can be useful for biomarking but is not neural connectivity. Statistical dependence measures should be specified: For example: ● Correlation ● Partial coherence ● Phase coupling ● Amplitude coupling ● Spectral Coherence ● Entropy ● Dynamic Causal Modelling (DCM) ● Granger Causality Any assumptions specified, for example: ● Linear vs unspecified ● Directional vs non-directional Any Cross-Frequency Coupling analysis should be explicitly specified, with the exact method for this and all parameters in detail. Niso et al (2022) further discuss that some mathematical understanding of the signal processing involved (in all EEG analyses) is required to analyse data. Specifically for connectivity metrics the problems associated with metrics sensitive to zero-phase lag coupling are not obvious from the outset.

○

1.5 Stimuli

We recommend that the researcher follows the suggested general checklist for stimuli and stimuli-related task description as detailed in Table 6, Section 2.6 in *General guideline for psychophysiological measures,* COMFOCUS D6.1 Guideline for measuring psychophysiological responses (link).

It should be noted that all stimuli features will have a possible effect on EEG signal, and they should ideally be standardised for factors not being researched, such as luminance, size, position, colour, shape, salience, clarity, complexity and quality/resolution (for visual), volume, clarity and quality (for audio).

Eye movements and blinking also affect EEG signal, so stimuli that cause a blink (e.g. sudden change in luminance) will confound the data with a large artifact, and stimuli which require/cause large eye movement will also cause confounds in the data. Within laboratory experiments, making stimuli small and central enough that participants do not need to move eyes will control for eye movement, preferably with a central fixation cross to fixate on. See design section (1.2.1).

1.6 Minimum reporting checklist

In addition to the general checklist suggested in Section 7.1 *Minimum reporting checklists for all studies using psychophysiological measures* in COMFOCUS D6.1 Guideline for measuring psychophysiological responses (link), we recommend the minimum reporting guidelines as detailed below.

In general, using the standard **Brain Imaging Data Structure (BIDS)** folder, data and metadata structure would be highly recommended for openness and ease of use of data. For information, see EEG-BIDS paper (Pernet et al, 2019), and online documentation (https://bids-specification.readthedocs.io/en/stable/04-modality-specific-files/03-electroencephalography.html). The metadata files required for BIDS-EEG are a) sidecar JSON, b) channels description, c) electrodes description, d) coordinate system, e) landmark photos. Your data can be checked using the BIDS Validator (https://bids-standard.github.io/bids-starter-kit/validator.html) to see if it meets the requirements. The advantage to this is that this data is now well-annotated for re-use and can be used with "off-the-shelf" toolboxes by both yourself and other researchers.

Furthermore, the Organisation for Human Brain Mapping (OHBM) Committees on Best Practices in Data Analysis and Sharing (COBIDAS) has published guidelines for EEG termed **COBIDAS-EEG** (Pernet et al, 2020), which includes recommended minimum reporting information for publication. If you cannot access the paper, the author's version of the paper is available on the OHBM website (https://www.humanbrainmapping.org/files/2020/COBIDAS-MEEG_NatNeuro_revised.pdf)

Created by researchers within the COBIDAS group, **ARTEM-IS (Agreed Reporting Template for EEG Methodology - International Standard:** https://artemis.incf.org/; https://github.com/INCF/artem-is) is also a template for minimal reporting (see Niso et al 2022, and Styles et al 2021), which has a web app to help researchers. It covers many points of minimal reporting discussed in this document, and can be used to generate templates to use for your study.

Table 6. Minimum reporting checklist for studies using Electroencephalography (EEG)

Category	Example
Data Structure & Description	Use Brain Imaging Data Structure (BIDS) for EEG (BIDS-EEG) structure for folders, data and metadata to increase reproducibility and ease analysis. BIDS validator can be used to check if data is well-annotated.
Participant Selection **Experimental Task Setup** **Experimental Task Information** **Task-free recordings** **Behavioural measures**	See Table 1: Recommendations for basic experimental attributes, from COBIDAS-EEG paper (Pernet et al, 2020, also see https://osf.io/a8dhx/ appendix tables) Split into what needs to be reported (23 items) and what needs to be in supplementary materials (3 items) In addition, report: ● Basic hardware, software and acquisition parameters: ○ e.g. Analog (in-built to amplifier) filter bandwidth (high pass, low pass) often not reported. ● EEG reference electrodes and impedances
Sensor Removal **Artifact removal**	See Table 2: Overview of data preprocessing steps, parameters that should be reported and their impact on reproducibility, from COBIDAS-

Physiological artifact removal **Downsampling** **Detrending** **Filtering** **Segmentation** **Baseline correction** **Re-referencing** **Normalization (for multivariate analyses)** **Spectral transformation**	EEG paper (Pernet et al, 2020, also see https://osf.io/a8dhx/ appendix tables) Split into parameters that need to be reported (24 parameters) and their impact (10 items)
Analysis **Network estimation** **Network metrics**	See Table 3: Necessary parameters to report in MEEG connectivity modeling to ensure reproduction of the method used, from COBIDAS-EEG paper (Pernet et al, 2020, also see https://osf.io/a8dhx/ appendix tables) 12 Parameters/categories are reported within these headings. In addition, report: ● Latency window to quantify Event-related potentials should be explicitly mentioned ● Spectral analyses should report boundaries of different frequency bands.

References

Barraza P, Dumas G, Liu H, Blanco-Gomez G, van den Heuvel MI, Baart M, Pérez A. (2019) Implementing EEG hyperscanning setups. *MethodsX.* Feb 26;6:428-436. doi: https://doi.org/10.1016/j.mex.2019.02.021

Choy, T., Baker, E. & Stavropoulos, K. (2022). Systemic Racism in EEG Research: Considerations and Potential Solutions. *Affec Sci* **3**, 14–20. https://doi.org/10.1007/s42761-021-00050-0

Clayson, P.E. Keil, A., Larson, M.J. (2022). Open science in human electrophysiology, *International Journal of Psychophysiology*, Volume 174, Pages 43-46, https://doi.org/10.1016/j.ijpsycho.2022.02.002.

Cooper, P.S., Baillet, S., Maroun, R.E.K., Chong, T, T-J. (2021). Over the rainbow: Guidelines for meaningful use of colour maps in neurophysiology. *NeuroImage*, Volume 245, https://doi.org/10.1016/j.neuroimage.2021.118628

Etienne, A. *et al*.. Novel Electrodes for Reliable EEG Recordings on Coarse and Curly Hair. (2020). *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, pp. 6151-6154. https://doi.org/10.1109/EMBC44109.2020.9176067

Farrens, J. L., Simmons, A. M., Luck, S. J., & Kappenman, E. S. (2019). Electroencephalogram (EEG) Recording Protocol for Cognitive and Affective Human Neuroscience Research. Protocol Exchange. https://protocolexchange.researchsquare.com/article/pex-779/v4. See also: https://static1.squarespace.com/static/5abefa62d274cb16de90e935/t/5df7db5956ec9170a9b402e1/1576524637645/Electrode_Application_Protocol_Final_Compressed.pdf

Franconeri, S. L., Padilla, L. M., Shah, P., Zacks, J. M., & Hullman, J. (2021). The Science of Visual Data Communication: What Works. *Psychological Science in the Public Interest*, *22*(3), 110–161. https://doi.org/10.1177/15291006211051956

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*, Vol 162, Pages 69-78, https://doi.org/10.1016/j.ijpsycho.2021.02.005.

Henrich, J., Heine, S. & Norenzayan, A. (2010) Most people are not WEIRD. *Nature* 466, 29. https://doi.org/10.1038/466029a

Kane, N., Acharya, J., Beniczky, S., Caboclo, L., Finnigan, S., Kaplan, P.W., Shibasaki, H., Pressler, R., van Putten, M.J.A.M. (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. *Clinical Neurophysiology Practice*, Volume 2, Pages 170-185, https://doi.org/10.1016/j.cnp.2017.07.002

Metzen, D., Genç, E., Getzmann, S. *et al.* (2022) Frontal and parietal EEG alpha asymmetry: a large-scale investigation of short-term reliability on distinct EEG systems. *Brain Struct Funct* **227**, 725–740. https://doi.org/10.1007/s00429-021-02399-1

Niso, G., Krol, L.R., Combrisson, E., Dubarry, A.S., et al. (2022) Good scientific practice in EEG and MEG research: Progress and perspectives. *NeuroImage,* Volume 257, https://doi.org/10.1016/j.neuroimage.2022.119056

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Pavlov YG, Adamian N, Appelhoff S, Arvaneh M, Benwell CSY, et al.(2021). #EEGManyLabs: Investigating the replicability of influential EEG experiments. *Cortex.* 144:213-229. doi: https://doi.org/10.1016/j.cortex.2021.03.013

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

Pernet, C., Garrido, M.I., Gramfort, A. et al. (2020) Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research. *Nat Neurosci* 23, 1473–1483. https://doi.org/10.1038/s41593-020-00709-0; see also https://cobidasmeeg.wordpress.com/ and https://osf.io/a8dhx/

Rae CL, Farley M, Jeffery KJ, Urai AE. Climate crisis and ecological emergency: Why they concern (neuro)scientists, and what we can do. *Brain Neurosci Adv*. 2022 Feb 28;6:23982128221075430. doi: https://doi.org/10.1177/23982128221075430

Smith, E.E., Reznik, S.J., Stewart, J.L., Allen, J.J.B. (2017) Assessing and conceptualizing frontal EEG asymmetry: An updated primer on recording, processing, analyzing, and interpreting frontal alpha asymmetry. *International Journal of Psychophysiology*, Volume 111, Pages 98-114, https://doi.org/10.1016/j.ijpsycho.2016.11.005.

Sinnott-Armstrong, W. & Simmons, C. (2021). Some common fallacies in arguments from M/EEG data. *NeuroImage,* Volume 245, 118725, https://doi.org/10.1016/j.neuroimage.2021.118725.

Styles, S.J., Kovic, V., Ke, H., Š o š ki ć, A., 2021. Towards ARTEM-IS: an evidence-based agreed reporting template for EEG methodology - international standard. *Neuroimage* 245, 118721. doi: https://doi.org/10.1016/j.neuroimage.2021.118721

Webb, E.K., Etter, J.A. & Kwasa, J.A. Addressing racial and phenotypic bias in human neuroscience methods. *Nat Neurosci* **25**, 410–414 (2022). https://doi.org/10.1038/s41593-022-01046-0

Yuval-Greenberg, S., Tomer, O., Keren, A.S., Nelken, I., Deouell, L.Y., (2008). Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron* 58 (3), 429–441. https://doi.org/10.1016/j.neuron.2008.03.027