Capturing the nature of events and event context using Hierarchical Event Descriptors (HED)

Kay Robbins[1], Dung Truong[2], Stefan Appelhoff[3], Arnaud Delorme[2,4], Scott Makeig[2]

[1]Department of Computer Science, University of Texas San Antonio San Antonio, Texas, USA

[2]Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California San Diego, La Jolla, California, USA 92903-0559

[3]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

[4]Paul Sabatier University in Toulouse, Toulouse, France

Corresponding author: Kay Robbins
Department of Computer Science
University of Texas at San Antonio
One UTSA Circle
San Antonio, TX 78249
Kay.Robbins@utsa.edu

Abstract

Because of the central role that event-related data analysis plays in EEG and MEG (MEEG) experiments, choices about which events to report and how to annotate their full natures can significantly influence the value, reliability, and reproducibility of MEEG datasets for further analysis and meta/mega-analysis. Current, more powerful annotation strategies combine robust event description with details of experiment design and metadata in a human-readable as well as machine-actionable form, making event annotation relevant to the full range of neuroimaging and other time series data. This paper dissects the event design and annotation process using as a case study the well-known multi-subject, multimodal dataset of Wakeman and Henson (*openneuro.org, ds000117*) shared by its authors using Brain Imaging Data Structure (BIDS) formatting (*bids.neuroimaging.io*). We propose a set of best practices and guidelines for event handling in MEEG research, examine the impact of various design decisions, and provide a working template for organizing events in MEEG and other neuroimaging data. We demonstrate how annotations using the new third-generation formulation of the Hierarchical Event Descriptors (HED-3G) framework and tools (*hedtags.org*) can document events occurring during neuroimaging experiments and their interrelationships, providing machine-actionable annotation enabling automated both within- and across-study comparisons and analysis, and describe a more complete BIDS formatted, HED-3G annotated edition of the MEEG sessions of the Wakeman and Henson dataset (*openneuro.org, ds003645*).

Keywords: Events, event annotation; Hierarchical Event Descriptors; HED; Brain Imaging Data Structure; BIDS; EEG; MEG; HED-3G; time series

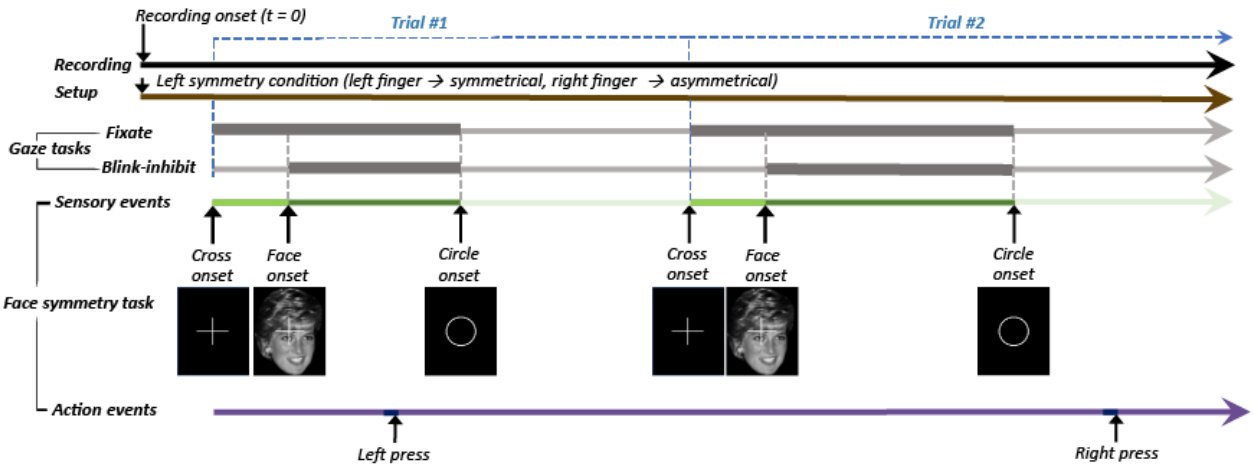
1. Introduction

EEG (electroencephalography) and MEG (magnetoencephalography) neuroimaging, collectively known as MEEG, are non-invasive brain imaging technologies for capturing neuroelectromagnetic brain dynamic records at millisecond-scale sampling rates. As MEEG records brain signals occurring on the time scale of thoughts and actions, event-related data analysis plays a central role in MEEG experiments. Good choices in how events are reported and annotated can greatly improve the immediate and long-term utility of the collected data. These annotation capabilities can be used to incorporate useful information about experimental design, experimental tasks, and other metadata with the data, as well as marking features discovered in the data itself, thereby enabling automated within- and across-study analyses.

An *event* is an identifiable process or condition that can be associated with a defined time period and distinguished from preceding, concurrent, and succeeding events or conditions. Events may also mark the occurrence of a time-delimited phase of or change in an ongoing process or condition. This paper introduces a practical event design strategy and illustrates a set of best practices for event reporting and annotation using a case study of a publicly-available multi-participant, multi-modal neuroimaging dataset from an experiment by Daniel Wakeman and Richard Henson (Wakeman & Henson, 2015), which will be abbreviated in the paper as W-H. The implementation uses the new third-generation formulation of the Hierarchical Event Descriptor annotation framework (HED-3G) with particular attention to the Brain Imaging Data Structure (BIDS) format (Gorgolewski et al., 2016) (Niso et al., 2018) (Pernet et al., 2019) (Holdgraf et al., 2019) in which the W-H dataset has been made available. The W-H dataset was initially released on OpenfMRI (*openfmri.org*) and then on its successor OpenNeuro (*openneuro.org*), both with access ID *ds000117*. We have recently shared on OpenNeuro a version of the W-H MEEG data (*ds003645*) with the more complete event record and annotation discussed in this paper.

The W-H experiment used here as a case study was conducted to develop methods for integrating multiple imaging modalities into analysis and for increasing the accuracy of functional and structural connectivity analyses. Nineteen participants completed two recording sessions spaced three months apart – one session recorded fMRI data (W-H-fMRI) and the other, simultaneously recorded MEG and EEG data (W-H-MEEG). During each session, participants performed the same simple perceptual task, responding to presented photographs of famous, unfamiliar, and scrambled faces by pressing one of two keyboard keys to indicate a subjective yes or no decision as to the relative spatial symmetry of the viewed image. Famous faces were feature-matched to unfamiliar faces; half the faces were female. The two sessions (MEEG, fMRI) had different organizations of event timing and presentation because of technological requirements of the respective imaging modalities. Each individual face was presented twice during the session. For half of the presented faces, the second presentation followed immediately after the first. For the other half, the second presentation was delayed by 5-15 face presentations.

Fig. 1 shows a schematic view of the unfolding of a typical run in the W-H experiment. All of the runs were conducted using the same equipment and positional seating of the participant, and this remained constant throughout the run (top black timeline). The bottom two timelines show the sensory events and participant actions (green and purple timelines, respectively).



Excerpt from a W-H event file based on this timeline.

onset	duration	sample	event_type	face_type	repetition_type	trigger	stim_file
23.400	n/a	25740	show_cross	n/a	n/a	n/a	cross.bmp
23.870	n/a	26275	show_face	famous_face	first_show	7	f032.bmp
24.081	n/a	26488	press_left	n/a	n/a	256	n/a
24.750	n/a	27225	show_circle	n/a	n/a	0	circle.bmp
26.457	n/a	29095	show_cross	n/a	n/a	n/a	cross.bmp
26.940	n/a	29634	show_face	famous_face	immediate_repeat	8	f032.bmp
27.913	n/a	30701	show_circle	n/a	n/a	0	circle.bmp
27.990	n/a	30789	press_right	n/a	n/a	4096	n/a

Figure 1. Schematic diagram of temporal organization and excerpt from a sample events file for a W-H recording. Upper left: Recording begins. Setup includes selection of the key assignment for responses in the face symmetry judgment task. The participant was asked to fixate on a central cross and to refrain from blinking while the face was being presented. Lower timelines: Sensory events were (sudden onset) visual image presentations; participant responses were key presses. Note: normal practice and lab equipment only records the offset of a physiological key press action; the key press trigger signal is preceded by hand electromyographic activation, finger pressure increase and movement, culminating in the electrical contact of the pressed key. Bottom table: a sample events file excerpt corresponding to the schematic timeline. (The events file format is explained in more detail in Section 3.1.)

Some of the participants were instructed to press a key with the left index finger to indicate above average facial symmetry (left symmetry key assignment), while other participants were instructed to press a key with the right index finger to indicate above average facial symmetry (right symmetry key assignment). The assignment was in effect for all of the recordings associated with a particular participant (brown timeline). The participants were instructed to fixate on the white cross and not to blink while the face was presented (thick gray gaze task timelines).

The W-H experiment has five main event marker types, which we will refer to as `show_cross`, `show_face`, `show_circle`, `left_press`, and `right_press`, in addition to some initial setup at the beginning of the recording. The fundamental problem addressed here is how to effectively annotate this information in a standardized form that is human-readable, machine-actionable, and analysis-ready without placing undue burden on the annotator.

The exploration begins with a brief introduction to the HED-3G system and explains the annotation process using the W-H experiment with a particular event design as a concrete example. Section 3 shows how these annotations are actually organized in a BIDS dataset to achieve machine-actionable, analysis-ready annotation. Using the example developed in Sections 2 and 3, Section 4 examines the event design and annotation process in more detail and proposes a set of guidelines for effective event design and annotation in MEEG. We discuss which events should be reported, how these events should be encoded, and the relationship of the encoded events to the task, experimental design, and the underlying purpose of the experiment. Section 4 also summarizes the importance and potential impact of good event design and best practices annotation strategies in making data more usable and the research based on it interpretable and reproducible. A brief review and roadmap for future HED development is given in Section 5.

2. Machine-actionable event annotation using HED-3G

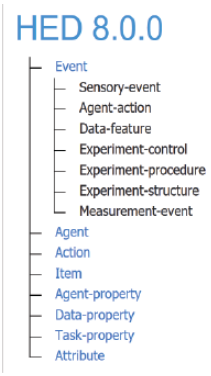
The HED system is based on a collection of hierarchical vocabulary structures (the base HED schema) in which each node is a defined term for describing events, condition variables, tasks, metadata, or the recording's temporal structure. HED was specifically designed to encode information in a both human- and machine-actionable format, specifically to enable validation, search, identification, and analysis of events in neuroimaging or other datasets that include time-marked events.

While the original HED implementation (first generation) focused mainly on a description of stimuli and responses (Bigdely-Shamlo et al., 2013). The second generation HED framework (Bigdely-Shamlo et al., 2016) developed for a closed group database project, included many vocabulary improvements and tools for validation and analysis. HED has recently undergone an extensive third-generation redesign (HED-3G) to enable capture not only basic event descriptions, but also experimental conditions, temporal structure, and event context (Robbins et al., 2020). HED-3G provides a readily extensible basis for easily interpretable annotation of time series datasets for use in analysis, re-analysis, and shared data mega-analysis. HED was accepted in 2019 as an (optional) standard for event annotation in BIDS formatted data. The current HED-3G development is aimed to ready HED for widespread use for data archiving, sharing, analysis, and mega-analysis.

2.1 A starting point for HED dataset event annotation

The HED-3G base schema has eight top-level or root nodes as shown by the partially expanded schema tree shown on the left in Figure 2. The very basic HED event annotation shown in the table inset on the right is our starting point for development of comprehensive annotation.

A rudimentary HED annotation of five W-H event types.



Event	HED
`show_face`	*Sensory-event, Visual, Experimental-stimulus, Face*
`show_circle`	*Sensory-event, Visual, Cue, Circle*
`show_cross`	*Sensory-event, Visual, Cue, Cross*
`left_press`	*Agent-action, Participant-response, Press, (Keyboard-key, Left-side)*
`right_press`	*Agent-action, Participant-response, Press, (Keyboard-key, Right-side)*

Figure 2. Left: partially expanded HED-3G schema tree. Right: basic annotation of the five main W-H event categories using HED-3G.

To annotate events, users create comma-separated lists of terms selected from the HED base schema to describe the events. In some sense, the annotation process can be thought of as using keywords from a structured vocabulary to tag the events. More information can be added to a tag by grouping the tag with other modifier tags using parentheses, as for example *(Keyboard-key, Right-side)* in Fig. 2.

Generally, the first step in event annotation is to select a term from the *Event* subtree to give a basic categorization of the event. Additional tags can be added to provide more description, but this top-level event categorization tag is often a primary search key for isolating events of interest. For comparisons of MEEG *across* studies, more detailed annotation can add significant value to the data. Annotations to address 'Which fingers were used to press the left and right keys?', 'How large were the cross, face image, and circle?', 'Where were these images presented on the screen?', or 'Were the image presentation onsets sudden (tachistoscopic)?' are examples of the types of additional details that might be of interest.

2.2 Short and long form annotation

A critical usability innovation in HED-3G is the stipulation that each HED schema node name must be unique (i.e., must only appear in one place in the schema). As a result, an annotator can tag using just a single node (e.g., the end node *Circle* in the *Item* hierarchy), rather than spelling out the full hierarchical schema path string (e.g., *Item/Object/Geometric-object/2D-shape/Ellipse/Circle*) as in the case of HED-1G and HED-2G. Automated HED tools can then perform the mapping from such short-form tags to include their complete (long-form) paths as

needed when the data are to be validated or analyzed. The expanded long-form annotations allow tools collecting related events for analysis to find HED strings that belong to more general categories – for example, searching for HED strings containing the more general term *2D-shape*, not only the more specific *Circle*. This type of organization is particularly useful for gathering data epochs that are time locked to a variety of events across datasets with some feature or features in common and/or annotated in different levels of detail.

In this paper, all of the HED tag examples are given in short form for readability. Examples of different forms of tag expansion into long form are given in the Supplementary Table 2. HED tools are now available to translate strings and files containing HED tags between the forms.

2.3 Specifying annotation "concepts" using HED definitions

Fig. 2 (above) gives a very minimalist HED-3G annotation for the five main types of W-H experiment event markers described schematically in Fig. 1. This level of annotation allows analysts to isolate events of different types such as basic stimulus events and participant actions, but does not provide sufficient detail to support advanced analysis and cross-study comparisons. Further, the annotation treats each event as occurring instantaneously, but the image presentation events have distinct onsets, durations, and offsets, all of which are known to affect brain dynamics measured by MEEG or fMRI.

HED user definitions allow annotators to document the structure as laid out in Fig. 1. HED-3G allows users to "define" or "declare" a name of their choosing and associate it with any group of tags. Users can then use the declared name in place of the individual tags during annotation. Definitions have the benefit of allowing users to create tagging short-cuts. This is especially useful if some tag groups are frequently reused.

Definitions also make it easier to identify and successively refine "concepts" by adding more details or refining tags to note particular aspects. Definitions improve annotation organization similar to the way functions improve programming code organization when refactoring inline code. Most importantly, HED-3G user definitions play an integral role in assisting data authors in documenting experiment architecture, event temporal extent, and other aspects of the dataset.

Here is a simple HED user-definition (*Circle-only)* representing the display of a circle:

(Definition/Circle-only, (Visual, (White, Circle), (Center-of, Screen)))

This definition describes the visual presentation of a white circle in the center of the screen. Of course, this annotation could be made more detailed, for example giving the size of the circle, the thickness of its perimeter, and the color of the background; these details can be added at any point in the annotation process. Once defined, annotators can use *Def/Circle-only* in annotations in place

of the longer defined tag string, *(Visual, (White, Circle), (Center-of, Screen))*. Below, we focus on the use of definitions to annotate more of the temporal structure of the participant experience.

The green timeline of Fig. 1 shows the time courses of the sensory events in the W-H data. The bright green bar marks the pre-stimulus period during which a white cross is displayed, while the dark green bar marks the time during which the face image is displayed. The light green bar marks the period during which a white circle is displayed. The boundaries between these displays are marked by the `show_cross`, `show_face`, and `show_circle` event markers, respectively. Thus, in the W-H experiment the face stops being displayed when the circle image starts. In addition, two gaze tasks (represented by the thick gray timelines in Fig. 1) coincide with these events: participants were asked to maintain eye gaze fixation on the white cross while it was displayed and to inhibit eye blinks during display of the face images.

Table 1 shows a revision of the table inset of Fig. 2 using definitions including *Onset* and *Offset* tags to encode the actual timeline of Fig. 1. The definition names (which are defined elsewhere) are bolded for readability. (See Supplementary Table 1 for completed annotation.)

Table 1. HED annotations of W-H event marker types that capture the event time evolution.

Event	HED
`show_cross`	*Sensory-event, Cue, (Def/**Cross-only**, Onset), (Def/**Fixation-task**, Onset), (Def/**Trial**, Onset), (Def/**Circle-only**, Offset)*
`show_face`	*Sensory-event, Experimental-stimulus, (Def/**Face-image**, Onset), (Def/**Blink-inhibition-task**, Onset), (Def/**Cross-only**, Offset)*
`show_circle`	*Sensory-event, Cue, (Def/**Circle-only**, Onset), (Def/**Trial**, Offset), (Def/**Face-image**, Offset), (Def/**Blink-inhibition-task**, Offset), (Def/**Fixation-task**, Offset)*
`left_press`	*Agent-action, Participant-response, Def/**Press-left-finger***
`right_press`	*Agent-action, Participant-response, Def/**Press-right-finger***

When a definition name is grouped with an *Onset* tag such as the *Face-image* defined name in the annotation for `show_face`, that annotation represents the start of an event with temporal scope. *Face-image* is assumed to be in effect until the next event in which a *Def/Face-image* tag appears grouped with an *Offset* tag (the annotation for `show_circle`). Referring to the event file excerpt of Fig. 1, a `show_face` event occurs at time 23.870s, while the next `show_circle` event occurs at 24.750s, Thus, the face image "endures" for 24.750 − 23.870 = 0.880 s.

Similarly, all of the sensory event and task timelines are encoded in Table 1. The *Press-left-finger* and *Press-right-finger* definitions of Table 1 are not grouped with *Onset* or *Offset* tags because here only the time of key release was recorded; thus we can only model these participant actions as instantaneous, single time point events.

Definitions allow annotators to express the structure of the experiment using higher level term concepts rather than repeating a mass of tagging details. Here, by defining "overview" descriptions

of the relevant sensory presentations *(Face-image*, *Cross-only*, *Circle-only)*, actions *(Press-left-finger*, *Press-right-finger*), and tasks (*Blink-inhibition-task*, *Fixation-task*), the annotator can quickly document the experimental structure in a manner that should also allow future tools to automatically extract and visualize dataset timelines. Additional event details can easily be added into the definitions at a later stage. Note that HED definitions cannot be nested.

An additional note about naming conventions and nomenclature. While HED tags are case insensitive, by convention HED tags start with a capital letter. Also, HED-3G tags cannot contain blanks, and tags representing multiple words are hyphenated. We use a font change in this paper to distinguish HED tags from other terms and names. Terms that appear in the actual event files (e.g., `show_face` or `event_type`) are printed in fixed point type and use underbars as word separators. This choice distinguishes HED annotations from identifiers used in event files, and allows tools to map event identifiers into program variables or structure fields.

2.4 Event context and temporal events

Effects of *ongoing context* on event-related MEEG brain dynamics have long been known (Squires et al., 1977) although are not frequently studied. HED-3G tools can automatically insert context information in a *Context* tag group when the annotation for an event is assembled at analysis time. During analysis tag expansion, tools can insert the tags associated with an ongoing (enduring) event into the *Context* tag groups of events that occur while the enduring event is ongoing. For example, suppose a participant presses a key while a movie clip is playing. After creating a *PlayMovie* definition to describe the movie presentation, the researcher can annotate the event marking the start of the movie with *(Def/Play-movie, Onset)* and the event marking the end of the movie with *(Def/Play-movie, Offset)*.

Tags that the researcher uses to annotate the key press event do *not* need to include information about the ongoing movie presentation. During event search and analysis, HED context tools (under development) will be able to automatically insert the context information (e.g., here, that a particular movie was playing (and/or a particular movie event was unfolding) at the time of the key press) into the *Event-context* tag group for the key press event. The goal here is that future HED context tools will support studies of consequences of previous events on behavior and brain dynamics associated with subsequent events (simple example: in ‘oddball’ experiment designs, the effects of rare target stimulus presentations and behavioral responses on brain and behavioral responses to immediately ensuing events).

2.5 Annotating experimental design and condition variables

The event sequence and the annotations of the previous section define what happens during the experiment, but do not convey the purpose of the experiment or how it relates to the underlying experimental design. A goal of HED-3G is to provide convenient mechanisms for annotating this

information in sufficient detail that tools can automatically extract and use experimental design information during analysis. In this section we introduce the *Experimental-condition* tag and use it with definitions to encode the W-H experimental design.

The W-H experiment uses a 3 × 3 experimental design. The two factors (i.e., design variables) are the face type and the repetition type, each with three levels. The primary author analyses (Henson et al., 2011) (Wakeman & Henson, 2015) (Henson et al., 2019) focus on face type (with levels corresponding to the display of famous, unfamiliar, and scrambled faces, respectively). The authors computed event-related potentials (ERPs) and some frequency-based measures for MEEG responses to different types of faces with an underlying purpose of improving source localization by leveraging participant information obtained from multiple imaging modalities.

Each image was shown twice during a session. The repetition type factor (with levels corresponding to the first display, an immediate repeat display, and delayed repeat display of an image) encodes the relative position of an image presentation with its match. The delayed repeat display level indicates that an earlier presentation of this image occurred 5 to 15 trials before this one. The repetition type design variable supports study of effects of image novelty versus repetition in the W-H data.

The information about the experimental conditions in effect for a data run is typically conveyed in one of two ways: either by inserting additional informational columns or by inserting additional events (rows) in the events file. The columns contain coding of the experimental condition, while the rows mark the starting and stopping of that experimental condition. The representation choice is up to the annotator. In the event design of this case study, the face type and repetition type information are included as the `face_type` and `repetition_type` columns of the events file, respectively (as shown by the excerpt in Fig. 1 above). That choice makes sense in this case because the selection of levels of the two factors changes so frequently across face presentations.

Table 2 summarizes the 3×3 experimental design and demonstrates how the information can be encoded using HED tags. The factor names (column 1) correspond to event file column names `face_type` and `repetition_type`, respectively, since factor information is being encoded in the columns in the events file. The levels (`famous_face`, `unfamiliar_face`, `scrambled_face`) for design variable `face_type`, appear as values in the `face_type` column of the events file.

Table 2. Experimental conditions with HED term definitions for the W-H experiment.

Variable	Level	Annotation
`face_type` (by trial)	`famous_face`	**Description**: A face that should be recognized by the participants. **HED:** *(Definition/**Famous-face-cond**, (**Experimental-condition**, Label/**Face-type**, (Image, (Face, Famous))))*
	`unfamiliar_face`	**Description**: A face that should not be recognized by the participants. **HED:** *(Definition/**Unfamiliar-face-cond**, (**Experimental-condition**,*

		*Label/**Face-typ**e, (Image, (Face, Unfamiliar))))*
	`scrambled_face`	**Description**: A scrambled face image generated by taking face 2D FFT. **HED:** *(Definition/**Scrambled-face-cond**, (**Experimental-condition**, Label/**Face-type**, (Image, (Face, Disordered))))*
`repetition_type` (by trial)	`first_show`	**Description**: Factor level indicating the first display of this face. **HED:** *(Definition/**First-show-cond**, (**Experimental-condition**, Label/**Repetition-type**, (First-item, Repetition-number/1)))*
	`immediate_repeat`	**Description**: Factor level indicating this face was the same as previous. **HED:** *(Definition/**Immediate-repeat-cond**, (**Experimental-condition**, Label/**Repetition-type**, (Next-item, Repetition-number/2)))*
	`delayed_repeat`	**Description**: Factor level indicating face was seen 5 to 15 trials ago. **HED:** *(Definition/**Delayed-repeat-cond**, (**Experimental-condition**, Label/**Repetition-type**, (Later-item, Repetition-number/2)))*
`key_assignment` (by participant)	`left_sym`	**Description**: Left finger key press means above average symmetry. **HED:** *(Definition/**Left-sym-cond**, (**Experimental-condition**, Label/**Key-assignment**, ((Mouse-button, Left-side), Symmetrical), ((Mouse-button, Right-side), Asymmetrical)))*
	`right_sym`	**Description**: Right finger key press means above average symmetry. **HED:** *(Definition/**Right-sym-cond**, (**Experimental-condition**, Label/**Key-assignment**, ((Mouse-button, Right-side), (Behavioral-evidence, Symmetrical)), ((Mouse-button, Left-side), (Behavioral-evidence, Asymmetrical))))*

The recommended strategy for annotating the factors and their levels, is to first create for each level, a HED definition that includes the *Experimental-condition* tag. The name of the definition is interpreted programmatically as the variable level (e.g., *Definition/Famous-face-cond* ⇒ `famous_face`, Table 2). The definition should also include a *Label* tag to indicate the name of the factor to which that level belongs (e.g., *Label/Face-type* ⇒`face_type`, Table 2). These elements appear in boldface in Table 2 to emphasize their role in documenting the experimental design. Notice how the events file excerpt in Fig. 1 has a column called `face_type` (the factor) and that values in this column such as `famous_face` correspond to the factor levels. Section 3 discusses in more detail how the definitions in Table 2 can be used in conjunction with `events.tsv` files to fully document the experimental design in the actual BIDS dataset.

Table 2 also lists a factor called `key_assignment`, which was used in the experiment for bias control. The key assignment factor (with levels `left_sym` and `right_sym` in Table 2) encodes the assignment of which index finger key press indicates a more symmetric face. In the `left_sym` condition, participants press a key with the left-index finger to indicate greater than average facial symmetry and press a key with the right index finger to indicate less than average facial symmetry. The left-right key assignment was counterbalanced across participants.

Notice that `key_assignment` does not correspond to a column in the table of Fig. 1. Because the level of this variable is constant for the entire recording, it may not make sense to use an entire column to encode the setting of the constant key assignment level. Instead, this variable might be better encoded by inserting an event at the beginning of the recording marking the onset of this condition. Here we insert an event at the beginning of the events file with an `event_type` value of either `left_sym` or `right_sym` to encode key assignment and the latency of the recording's first data point. We also insert an additional event (with an `event_type` of `setup`) at the beginning of the recording to gather information common to all of the recordings in dataset.

HED tools now under development will be able to automatically extract the design matrix and other statistics about the experimental design from the definitions that include the tag *Experimental-condition* and the events associated with these definitions. The definitions only need the *Label* and *Experimental-condition* tags to enable tools to identify the experimental control information for each recording.

The next section explains how the annotations developed in this section are actually mapped into a BIDS dataset and addresses some practical issues of annotation.

3 HED annotation of a BIDS-formatted dataset

BIDS recommendations for archival data storage have quickly become a *de facto* standard for sharing raw neuroimaging data. This section demonstrates how HED-3G event annotations are actually mapped into machine-actionable annotation of datasets organized according to BIDS specifications. BIDS specifies a particular dataset directory structure, naming conventions, and permitted image data formats, making it easier for users and tool developers to access data without manual recoding. In BIDS-formatted datasets, much of the metadata is located in `.json` text files called "sidecars", and naming conventions associate the sidecar metadata with the data files.

Importantly, the metadata describing the meanings of the columns (and the HED annotations) in the individual scan events files are contained in BIDS .json event files. When the same metadata applies to many data files, BIDS allows metadata files to be placed higher in the dataset directory hierarchy; their definitions will then be inherited by event files in dataset sub-directories (the *BIDS Inheritance Principle*), thereby avoiding the need to repeat the same metadata within multiple files in lower levels of the BIDS folder hierarchy. The BIDS naming convention associates the `..._events.tsv` files with relevant metadata `..._events.json` sidecar files, here and most importantly, the top (full dataset-level) `..._events.json` file.

Table 3 below summarizes different mechanisms for including HED annotations in a BIDS dataset. This case study includes HED information ONLY in the top-level `..._events.json` sidecar file contained in the dataset root directory. That information is keyed to the column names of the individual `..._events.tsv` files (Fig. 1 and Table 4 below) located at the lowest level of the dataset, each containing the list of events in the corresponding recording.

Table 3. Mechanisms HED annotations in BIDS `.json` and `.tsv` metadata files.

BIDS folder level	Information file	Function
Dataset	`..._events.json`	Provides descriptions of the columns that are applicable to all of the scan-level **`...events.tsv`** files in the dataset. [The 'HED' keys in this JSON dictionary link HED annotations to values in the events files.]
	`participants.tsv`	Lists the participants. [A 'HED' column may be used to add *participant-specific* information in HED annotation.]
Subject	`sessions.tsv`	Lists the sessions per participant. [A 'HED' column may be used to add *session-specific* information in HED annotation.]
Session	`scans.tsv`	Lists the scans in the session. [A 'HED' column may be used to add *scan-specific* information in HED annotation.]
Modality (Scan)	`..._events.tsv`	Lists the events in the scan. The column meanings and associated HED tags are given in the dataset-level **`..._events.json`** file**.** [A 'HED' column in **`..._events.tsv`** gives *event-specific* information in HED annotation.]

As summarized in Table 3, it is also possible to incorporate HED annotations in other BIDS `.tsv` files by including an extra column called `HED`. These annotations are particular to the row of the file and should only contain HED strings with no definitions. For example, a HED string appearing in the `HED` column of `participants.tsv` pertains to the participant described in that row. In annotating more complex experiment designs, some HED information could be placed most efficiently in any or all of the four BIDS `.tsv` file types listed in Table 3 (if present) as well as in additional `..._events.tsv` sidecars placed at lower levels in the dataset hierarchy, possibilities that for simplicity we do not discuss further here.

Users of EEGLAB (*sccn.ucsd.edu/eeglab*) (Delorme & Makeig, 2004) will notice in Table 3 that the BIDS file level naming approach is rather fMRI-centric (versus MEEG-centric). In BIDS, a data recording is termed a "scan" and the entire collection of data from an experiment is called a "dataset". In contrast, an EEGLAB "dataset" is (typically) a single continuous data recording (i.e., a BIDS "scan"), and the collection of the entire data collection for an experiment is termed a "study".

3.1 BIDS events.tsv files

At the lowest, single scan level, BIDS events files are tab-separated value (TSV) formatted text files with file names ending in `_events.tsv` (prefixes in the name give subject information, etc.). The first line in a BIDS events file is a header line identifying each column, and each subsequent line corresponds to an event marker in the data. Table 4 shows the excerpt of the events file of Fig. 1, color-coded to indicate the source of the expanded event annotations of Table 6 (see below).

Table 4. Excerpt of the individual scan-level `events.tsv` file illustrated in Fig. 1, with color-coding to key its expansion in Table 6 (following).

onset	duration	sample	event_type	face_type	repetition_type	trigger	stim_file
23.400	n/a	25740	show_cross	n/a	n/a	n/a	cross.bmp
23.870	n/a	26275	show_face	famous_face	first_show	7	f032.bmp
24.081	n/a	26488	press_left	n/a	n/a	256	n/a
24.750	n/a	27225	show_circle	n/a	n/a	n/a	circle.bmp
26.457	n/a	29095	show_cross	n/a	n/a	n/a	cross.bmp
26.940	n/a	29634	show_face	famous_face	immediate_repeat	8	f032.bmp
27.913	n/a	30701	show_circle	n/a	n/a	n/a	circle.bmp
27.990	n/a	30789	press_right	n/a	n/a	4096	n/a

BIDS requires event files to have `onset` and `duration` as their first and second columns, giving the onset time and duration of each event in seconds, respectively. Users may add additional columns as needed. All columns in the event file should be documented in one or more accompanying JSON-format sidecar files as described in the next section. The shaded columns of

Table 4 also have HED annotations. Questions such as, 'Which columns should be reported?' and 'How should column values be encoded?' are addressed in Section 4.

BIDS events files have two types of columns: categorical and value. **Categorical columns** define a small number of distinct levels or categories, representing text or numeric values. Other columns are **value columns**. The HED string for one event in the `..._events.tsv` file for one W-H recording in Table 4 has categorical columns `event_type` (blue), `face_type` (plum), `repetition_type` (green), and `trigger`, each with a relatively small number of distinct levels. Value columns are `onset`, `duration`, and `sample`. The `stim_file` column (tan) could be treated either as a categorical or as a value column depending on the number of distinct stimulus images. Here we treat `stim_file` as a value column because of the relatively large number of images used in the W-H experiment. The distinction between categorical and value columns is important mainly because HED annotations are encoded differently for the two types of columns.

3.2 BIDS events.json sidecar files

HED annotations of many experiment datasets can use a common and relatively simple event-design strategy across recordings that requires only a single `events.json` file placed at the top level directory of the dataset to provide complete machine-actionable event annotation when combined with the values in the individual `events.tsv` files. In general, an organization using a single dataset-level `events.json` sidecar is easier to annotate, understand, and maintain so that is the organization we focus on here. The annotated W-H case study of Section 2 assumes that all of the annotation for the dataset events is in a single `events.json` sidecar file located in the top level dataset directory. Table 5 shows a portion of this single `events.json` sidecar file for the annotation of events in the W-H dataset. (Supplementary Table 1 shows the complete sidecar.)

Table 5. Excerpt of the top (dataset) level events file (`events.json`) for the W-H data. See also Supp. Table 1.

```
{
 "onset": {
    "Description": "Onset of the event in seconds relative to the start of the recording.",
    "Units": "s"
 },
 . . .

 "repetition_type": {
    "Description": "Design variable for the first or a delayed showing of the image.",
    "Levels": {
       "first_show": "Factor level indicating the first display of this face.",
       "immediate_repeat": "Factor level indicating this face was the same as previous one.",
       "delayed_repeat": "Factor level indicating this face was seen 5 to 15 trials ago."
     },
    "HED": {
       "first_show": "Def/First-show-cond",
       "immediate_repeat": "Def/Immediate-repeat-cond",
       "delayed_repeat": "Def/Delayed-repeat-cond"
     }
```

```
 },
 "stim_file": {
     "Description": "Path of the stimulus file in the stimuli directory.",
     "HED": "(Image, Pathname/#)"
 }
 . . .
}
```

The `events.json` sidecar files are structured as dictionaries. The excerpt shown in Table 5 has three top-level dictionary keys (`onset`, `repetition_type` and `stim_file`). These keys correspond to column names in the events file excerpt shown in Table 4. (Here the annotations for the `sample`, `event_type`, `face_type`, and `trigger` columns are omitted for readability; they are included in Supplementary Table 1.) HED tools associate column metadata with particular columns in the event file using these column names. Besides keys corresponding to column names, BIDS users may use additional top-level keys to include additional metadata in the JSON sidecars.

The values associated with the top-level keys in `events.json` sidecars are themselves metadata dictionaries. Any of these metadata dictionaries that contains the word `HED` as a key is considered to contain HED annotations and is processed by the HED validator during BIDS validation and by the HED tools preparing the data for analysis. All of the annotations of Tables 1 and 2 and corresponding definitions are also stored in the full `events.json` sidecar file (Supplementary Table 1).

In Table 5, the metadata dictionaries associated with `repetition_type` and `stim_file` have '`HED`' keys and hence are HED annotations. In contrast, the metadata dictionary associated with top-level key `onset` does not include a '`HED`' key, so it is considered to be an unannotated column and is ignored by HED. The value that the '`HED`' key indexes in a metadata dictionary determines the format of the HED annotation. If the '`HED`' key references a dictionary, HED assumes a categorical column, while if the '`HED`' key references a string, HED assumes a value column. In either case, HED uses the corresponding dictionary values to annotate the event.

Categorical column annotations in `events.json` sidecar files include a separate HED annotation for each categorical value that appears in the corresponding column of the `events.tsv` file (e.g., here `first_show`, etc.). Value column annotations use a single HED string with a hash symbol (#) value placeholder to annotate each column. When the complete annotation for an event is assembled, HED tools replace the hash symbol with the relevant `events.tsv` file value from the respective column and row. The next section explains how the annotation for an event is assembled by combining event information in the …`_events.tsv` files with the HED annotation information in the …`_events.json` sidecar dictionaries.

3.3 Assembling and using the complete event annotation

HED tools gather the BIDS `events.json` sidecars entries applicable to an events file and assemble a single HED string representing the annotation for each event (i.e., for each line in the event file).

The assembled HED string annotation for the second event (`show_face`) in Table 4 (above) is shown in Table 6. Parts of the HED string are color-coded to indicate which column annotation that portion corresponds to. The corresponding columns in the `events.tsv` file of Table 4 are displayed in the same respective colors.

Table 6. Assembled HED string for the first showing of an image of a famous face (row 3, Table 4), in which the W-H data also marks the end of the cross-only presentation, and the onset of a blink inhibition period.

Sensory-event, ***Experimental-stimulus***, (***Def/Face-image***, *Onset*), (***Def/Blink-inhibition-task***, *Onset*), (***Def/Cross-only***, *Offset*), ***Def/Famous-face-cond***, ***Def/First-show-cond***, (***Image,*** *Pathname/f032.bmp*)

To annotate the `show_face` event in the second row of the `events.tsv` file, HED tools look up the column annotation for each value in the accompanying `events.json` sidecar. The `onset` and `duration` columns are not annotated for this example (although they could have been treated as value columns) so they are skipped. The `show_face` value of column `event_type` is translated into its corresponding HED annotation (Table 1) and concatenated to the assembled annotation (light blue shading). Next, the annotation for `famous_face` in the `face_type` column is looked up in the `events.json` sidecar dictionary and appended (plum shading). Then the category `first_show` in the `repetition_type` column is looked up, and the corresponding HED annotation is included (light green shading). Finally, the string filename value in the `stim_file` column is substituted for the # in the corresponding annotation (tan shading). The `sample` and `trigger` column values are skipped in this process, because they have no HED keys in the sidecar `events.json` sidecar dictionary.

During analysis, HED tools may expand the definitions so that their values are available for searching and filtering. (Supplementary Table 2 shows several forms of the assembled annotation of Table 6 and demonstrates how the *Def-expand* tag is used with the substituted definitions to accomplish this expansion.) Combining the information in the BIDS `events.tsv` files with the appropriate `events.json` sidecar annotation file(s) enables powerful automated tools to be implemented. For example, from this information, a task list, the underlying experimental design, and the temporal structure of a recording could be automatically extracted and visualized. Extensive statistics about the number of events with different properties could also be computed. Data could be separated into event-locked epochs with similar HED tags fitting a simple or complex description, and automatically bootstrapped to look for differences associated with different experimental parameters. Complex searches could be conducted across datasets (including datasets using different tasks and experimental designs) without need for manual re-coding.

4. Best practices in event design and annotation

A myriad of events, overt or covert, planned or unplanned, may unfold during the execution of an experiment; it may not be possible to record them all, nor be feasible to describe precisely their every detail. We use the term ***event design*** to refer to the process of identifying, organizing, reporting, and sufficiently annotating the nature of recorded events to a degree essential to complete interpretation of the event-related MEEG dynamics recorded during the experiment. *These events and descriptions should include all that is relevant to both current, planned and future potentially fruitful analyses*. Event design should be the first step in augmenting a dataset with HED annotation. **How a researcher chooses to organize, report, and annotate events can completely change the capacity of a given dataset to support analysis, reuse, and reproducibility.** Event design should be the first step in augmenting the value of a dataset with HED annotation.

Thus, best practices for event design must encourage researchers to look beyond the immediate use of their data to broader questions. In particular: ***Which aspects are potentially important to future analysis*** (performed either by the data authors or others)? These analyses are likely to include meta-analyses and mega-analyses (Costafreda, 2012) (Boedhoe et al., 2019) (Bigdely-Shamlo et al., 2019) across shared datasets that may involve different designs, participant tasks, experimental conditions, and event types. Details of events not relevant to initially planned analyses might prove valuable to later deeper and/or broader analyses. As MEEG data sharing, re-analysis, and meta/mega-analysis becomes the norm, prolonged credit will no doubt accrue to data authors whose care in performing data annotation make their datasets able to contribute widely to analysis and discoveries even far into the future.

The event design process has two steps: first *identifying* which events to report and then *mapping* these events into usable annotations. **Ideally, the event design process should be performed before data collection begins**, as we find that the event design process clarifies what is being measured and whether those measurements can be used to achieve experimental goals. In any case, most of the information required by a good event design process will be required in publications reporting the work, so performing a preliminary event design can help to assure that important details are not confused or overlooked later. In this section, we discuss the event design process and suggest guidelines for it using the W-H dataset as a case study. When HED annotation is performed completely *post hoc*, beginning the annotation by going through the event design process is useful for deciding how to best annotate the data.

4.1 Event design for the W-H experiment

The W-H event design developed in Sections 2 and 3 above is not the one distributed with the original shared OpenNeuro dataset *ds000117*, but was developed by us based on the recommended event design practices with the generous assistance of the data authors Wakeman and Henson, to make additional event type and timing information available in the data. The MEEG data of the redesigned dataset are available as OpenNeuro dataset *ds003645*. Table 7 shows the beginning rows of the MEEG events file for the first run (scan) of participant 02 using the recommended event design approach.

Table 7. MEEG event file run 1 for participant 02 using the recommended event design.

onset	duration	sample	event_type	face_type	repetition_type	trigger	stim_file
0.00400	n/a	1.000	setup	n/a	n/a	n/a	n/a
0.00400	n/a	1.000	left_sym	n/a	n/a	n/a	n/a
24.20582	n/a	6051.455	show_face	unfamiliar_face	first_show	13	u032.bmp
25.03127	n/a	6257.818	show_circle	n/a	n/a	0	n/a
25.15400	n/a	6288.500	left_press	n/a	n/a	256	n/a
26.73127	n/a	6682.818	show_cross	n/a	n/a	n/a	n/a
27.24582	n/a	6811.455	show_face	unfamiliar_face	immediate_repeat	14	u032.bmp
27.89309	n/a	6973.273	left_press	n/a	n/a	256	n/a
28.09582	n/a	7023.955	show_circle	n/a	n/a	0	n/a
29.79582	n/a	7448.955	show_cross	n/a	n/a	n/a	n/a
30.35309	n/a	7588.273	show_face	unfamiliar_face	first_show	13	u088.bmp
31.18400	n/a	7796.000	show_circle	n/a	n/a	0	n/a
32.88400	n/a	8221.000	show_cross	n/a	n/a	n/a	n/a
33.36036	n/a	8340.091	show_face	unfamiliar_face	first_show	13	u084.bmp
34.36400	n/a	8591.000	show_circle	n/a	n/a	0	n/a
36.06400	n/a	9016.000	show_cross	n/a	n/a	n/a	n/a

The event design of Table 7 marks the onsets and offsets of all of the experimental sensory presentations and participant motor responses using the annotations and encoding of the `event_type` column of Table 1. Further, the design encodes the 3 × 3 experimental design (using information in the `face_type` and `repetition_type` columns and the encoding described in Table 2.) Table 7 adds two events (`setup` and `left_sym`) at the time of the first data sample. The `setup` is a convenience event (meta-event) used to capture annotations applicable to all recordings in the dataset, such as the visual presentation screen size and distance (when available). The `left_sym` meta-event is used to capture the bias-control key assignment factor (either `left_sym` or `right_sym` depending on the participant). Since this key assignment is in effect for the entire recording, it is more efficient and clearer for tools to encode it as a meta-event with an onset (at the time of the first data sample) rather than giving it its own column in the `events.tsv` files with the same value repeated for every response event.

The event design also includes the sample number of the event onset in column labeled `sample`. This column is recommended in the BIDS standard and is good practice since the precision of the onset values is left completely open and accurate event timing is extremely important for MEEG analysis. The `trigger` column is not necessary, because the information is already encoded in the `face_type` and `repetition_type` columns, but we have retained it to maintain the connection with the original shared dataset. Table 8 shows, for comparison, a sample of the event file for the MEEG portion of the W-H data as originally shared. The `events.tsv` files only give the onsets of the face presentations and contain no markers for other sensory presentation or participant responses, limiting the usability of the data for analysis, further analysis, and meta/mega-analysis.

Table 8. MEEG event file for run 1 of session 1 of subject 01, as originally shared.

onset	duration	onset_sample	stim_type	trigger	stim_file
24.2073	0	26628	Unfamiliar	13	meg/u032.bmp
27.2473	0	29972	Unfamiliar	14	meg/u032.bmp
30.3545	0	33390	Unfamiliar	13	meg/u088.bmp
33.3618	0	36698	Unfamiliar	13	meg/u084.bmp

Table 8 is considerably shorter and narrower than Table 7 (our recommended version), but is missing critical information (e.g., `repetition_type` and all the events marking presentations of the fixation cross and focusing circle, and the key press events). Difficulties introduced for downstream analysis by not recording and reporting all possible sensory and participant action events are discussed in more detail in Section 4.3 and Section 4.4, respectively.

Another difficulty in Table 8 is the use of non-orthogonal encoding of the experimental design in the event-recording hardware system `trigger` column, whose 12 distinct values are shown in Table 9.

Table 9. The 12 trigger values from the W-H (shaded rows) and their respective interpretations.

0	5	6	7	13	14
Cross only	Famous Initial	Famous Repeat	Famous Delayed	Unfamiliar Initial	Unfamiliar Repeat
15	**17**	**18**	**19**	**256**	**4096**
Unfamiliar Delayed	Scrambled Initial	Scrambled Repeat	Scrambled Delayed	Left press	Right Press

While it is possible to HED-tag each trigger value to indicate which factors and levels it represents, the non-orthogonal or mixed encoding makes downstream analysis much more likely to require manual re-coding, thereby making the dataset difficult to include in further analysis and mega-analysis. In the recommended design (Table 7) the independent factors `face_type` and the `repetition_type` are represented by independent columns in the events file, making it easy for automated processing to detect the 3 × 3 design. Encoding of experimental conditions is discussed in more detail in Section 4.5.

4.2 Pitfalls in reporting events by-trial

An overall guideline for choosing event markers strongly favors expressing each relevant event with its own event marker in the event file. Where relevant, the marked events should also include offsets of all processes having appreciable duration, as well as their onsets. In some cases, intermediate events may be important for analysis, for example onsets of individual syllables in spoken words or critical points of hand/arm motion capture dynamics in reach trajectories.

Guideline 1: Event files should be organized by event. *Event files should always report one event per line. When computation of response times or delays are needed or convenient, the event file should still include rows in the events file representing the onsets and offsets of the actual events used to compute these response times or delays. These reported events should include all possibly relevant onsets and offsets of relevant sensory stimuli, motor actions, participant tasks and task conditions, condition changes during the recording, time organization, plus setup meta-event information organized during event design.*

While this recommended *by-event* organization may seem logical, many deposited datasets in fact instead use a *by-trial* organization or some hybrid. *By-trial* organization treats each trial as a single event (given one row in the event file) and expresses all other relevant events as offsets from the trial latency in the data. Such *by-trial* organization has many disadvantages for event-related analysis, most prominently a lack of clarity with respect to the timing of other most likely MEEG-affecting events. As an illustration, consider the sample of an event file originally shared for the fMRI portion of the W-H experiment, as shown in Table 10.

Table 10. The fMRI event file run 1 of session 1 for subject 01 as originally shared.

onset	duration	cross_duration*	stim_type	trigger	button_pushed	response_time	stim_file
0	0.908	0.534	FAMOUS	5	4	2.158	func/f013.bmp
3.273	0.962	0.586	FAMOUS	6	4	1.233	func/f013.bmp
6.647	0.825	0.546	UNFAMILIAR	13	4	1.183	func/u014.bmp

*Note this column, mistakenly labeled `circle_duration` in the original distribution, has been corrected.

When motor response events are reported only as `response_time` delays, it is not always clear whether the time is relative to the trial anchor event or to some other event. Here, events that occur before the anchor event are not always expressed with a negative delay (e.g., here `cross_duration` is positive, though the cross display occurs before the anchor face presentation onset event). Which actual event is the anchor event corresponding to the trial `onset` is not indicated and it may not be clear which actual event within the trial actually corresponds to the `onset` column. While it is possible to calculate the onsets and offsets of the visual stimuli from the various durations and response times, a data user would have to do a very careful analysis of the documentation and published papers to correctly identify the sensory and motor action event

onsets and offsets. Performing this anew for each shared dataset in some future mega-analysis across shared datasets would be either infeasible (or at best heroic).

By clearly identifying *all* of the experiment sensory events in a column named `event_type` or something similar, the design of Table 7 makes processing much easier. To reiterate, *identifying all event onsets and offsets* is increasingly important for many analyses, in particular those that use standard or new methods to model the complex, interacting effects of events on cognition and MEEG dynamics.

A second issue with *by-trial* organization of an event table is lack of extensibility. For consistency, each row in *by-trial* reporting contains information about the event sequence that is consistent across trials. Often, condition changes and other events need to be recorded outside the strict *by-trial* structure, thereby complicating the annotation process. When later adding event markers to the event file identifying additional events in the data (such as blinks, alpha spindles, interictal spikes, background noise outbreaks, etc.), researchers must decide whether to add additional columns and express the new times in terms of trial offsets or to add additional rows and treat the new markers as separate non-trial events. The difficulty with the latter approach is that the marked event times are likely to cross trial boundaries, thus requiring dataset-specific manual coding and analysis to unwind the information about those events. Operations such as regressing out the effects of overlapping events or determining effects of ongoing event context cannot be performed without first obtaining a distinct, well-ordered record of the dataset event onsets and offsets.

4.3 Documenting sensory presentations

Guideline 2: All known sensory presentations that are intended to or may affect the MEEG data should be reported and annotated. *Sensory presentations (including onsets, offsets, transitions between trial presentations and other known or easily computed significant moments) should be documented as events. In addition to the formally designated experiment "stimuli," dataset sensory presentations may include instruction delivery, sensory or other feedback, presentation of auxiliary stimuli such as fixation points or filler images to mark time, auditory cues, changes in background, and unplanned events noted as having occurred in the environment. The role of each sensory presentation within the task and experiment as well as a description of the sensory presentation itself should be documented. Event annotation should aim to document* ***what the participant experiences.*** *At a minimum, thoughtfully detailed reporting of participant sensory experience allows analysts to regress out the influences of other sensory presentations on dynamics associated with presentations of the primary stimuli; nonlinear modes of analysis may benefit still more from this information, quite possibly in ways currently undocumented.*

As first shared, the shared W-H MEEG dataset noted only face image presentation onsets, while the W-H fMRI dataset also included cross duration and key press response times as well as indicating which key was pressed (left or right). Papers published by the authors on this dataset included a somewhat more complete description of the event sequence. Table 11 summarizes the documented sensory presentations for the W-H dataset to indicate how event design for this experiment might proceed.

Table 11. Sensory presentations with HED term definitions for the W-H experiment.

Name	Concept description	HED definition body
Face-image	A happy or neutral face in frontal or three-quarters frontal pose with long hair cropped presented as an achromatic foreground image on a black background with a white fixation cross superposed.	*Visual, (Foreground-view, ((Image, Face, Hair), Color/Grayscale), ((White, Cross), (Center-of, Screen))), (Background-view, Black)*
Cross-only	A white fixation cross on a black background in the center of the screen.	*Visual, (Foreground-view, (White, Cross), (Center-of, Screen)), (Background-view, Black)*
Circle-only	A white circle on a black background in the center of the screen.	*Visual, (Foreground-view, ((White, Circle), (Center-of, Screen))), (Background-view, Black)*

Note that the *Face-image* definition focuses on the basic description of the presentation. In other words, the description of the sensory presentation focuses on annotation from the perspective of what the participant experiences (here, a sudden face image presentation on a black screen). The definition does not include information about the face category (famous, unfamiliar, or scrambled) that is dependent on the experiment design, nor about subject pose and gender. This information is specific to the individual face, while face image category information has an explicit role of the experimental design.

Notice also that none of these HED *definitions* contain tags from the HED schema *Event* subtree (e.g., *Event/Sensory-event* or *Event/Agent-action*). The HED schema *Event* tag and its child node tags are used to annotate the general type of an event as associated with markers in the events files. In general, **HED *Event* tag annotations should always appear outside of concept definitions** since multiple definitions are likely to be used in the annotation of each event while, ideally, a single *Event*-tree tag defines the overall category of the event markers. Following this rule also simplifies searching and filtering operations in downstream analysis.

Annotation during acquisition. We found some ambiguity in the published description of the W-H MEEG experiments. When did the first trial begin? Did recording begin at the start of the first trial? If not, was a white circle displayed at the beginning of the recording? To avoid such ambiguities, it is a good practice to write experimental control scripts that automatically output

event markers for every sensory presentation event as well as the data itself. We will be contacting experiment control application developers to encourage that this capability be built in, as real-time annotation of at least machine-controlled events can be far more efficient than *post hoc* hand coding.

4.4 Documenting participant responses

Guideline 3: Instructed participant motor responses (and any other recorded participant actions) should be reported as events. *Instructed participant responses or actions should be marked as individual events (or event sequences) rather than reported only as reaction times and/or by noting the category of the participant response (e.g., for the W-H experiment, as 'symmetric' or 'asymmetric'). Motor actions themselves, their planning, and accompanying and ensuing assessment processes are all supported by brain dynamics that are very likely to be reflected (in part) in MEEG data features. As with sensory presentations, motor responses should be first annotated from the perspective of* ***what the participant does****, not what it means in terms of the experiment design and task. At a minimum, the annotation should document who acts and what action they take. The experiment control program's handling of correct, incorrect, and omitted response actions should also be articulated, particularly if these affect the nature of later stimuli. Other types of participant actions, instructed or incidental, may also be documented. If these actions were not instructed, they are not likely to be part of the initial experiment design, so they may need to be entered as data features post hoc..*

In the W-H experiment, participants were instructed to press one of two keys with their respective left and right index fingers to indicate their assessment of facial symmetry. The symmetry evaluation task was unrelated to the experimenters' own true objective in running the experiment. Perhaps for this reason, the participant responses were not fully documented in the W-H data as originally shared, and there was no indication in the dataset documentation of what would occur when or if the participant withheld a key press entirely. Thinking more broadly about potential further uses for the data (e.g., when building the event-design) may hopefully inspire data authors annotating their data to make it fit for a broader range of uses and sharing, thereby considering it worthy to add all available detail about subject performance to the shared dataset to enhance continued dataset usability. Here, for example, the W-H face symmetry evaluation task might itself be of some future interest, as might also be how the pose or gender of a presented face affected brain dynamics and motor responses. Such readily recorded variables might also be treated as dependent variables to strengthen the statistical reliability of effects of interest in any analysis of the data. Both researchers and research funders are gradually realizing that the continuous increase in efficient and available computing power can make much research focused on innovative analysis of existing data much more affordable than funding a plethora of relatively

small, single-use data collection projects as has been the pattern in the past, a development that is also forcing changes in the education of neuroimaging scientists.

Note again that typical recordings involving key presses actually record the *culmination* of the physiological pressing action, not its physiological muscle activation or brain signal precursors. The ensuing process of key release may also be reflected in the recorded brain dynamics; however, none of this behavioral and psychophysiological information is typically recorded, neither in past nor still current practice. Growing appreciation of the importance of the embodied cognition perspective on mental life (Shapiro, 2019), new lightweight, low cost methods of recording details of brain activities and motor behavior of experiment participants (Casson, 2019) (Jas et al., 2021) (Vitali & Perkins, 2020), and emergence of the practice of recording both brain activity and behavior (as well as psychophysiology) at higher resolution (sometimes termed Mobile Brain/Body Imaging or MoBI) (Makeig et al., 2009), make development of a suitable data annotation framework ever more urgent.

4.5 Documenting experimental conditions, controls, and designs

Guideline 4: Experimental conditions, both fixed and changing, should be identified, whether they are part of the experimental design or are put in place to control experimental bias. *All experimental conditions should be documented, not just the main design variables. Full documentation allows researchers to systematically test for statistical differences in data features under various conditions. The explicitly stated experimental design provides the obvious factors to be annotated.* ***However, any aspect of the experiment that was controlled for bias can provide a condition for annotation.*** *Elements that are counterbalanced or randomized in a specified range should always be given serious consideration for explicit annotation as experimental conditions. The span of each condition should also be identified. Was the condition varied by trial, by block, by run, by session, or by participant? If so, how and when – precisely?*

Table 2 lists the experimental condition types or factors along with the individual levels or factor values that we have identified for the W-H dataset, with proposed HED definitions for each factor. Of the three experimental condition types listed in Table 2, `face_type` and `repetition_type` correspond to the original 3×3 experimental design, while participant-wise assignment of `key_assignment` was introduced to control bias. Notice that each definition in Table 2 includes both a term definition and a label. The label name corresponds to the design variable, while each definition name corresponds to a level or value of that variable. Annotation of condition variables allows tools to automatically extract from the event information the experiment design matrix, experiment timeline, and other experiment statistics.

For the W-H dataset, another possible condition type is face image sex (with levels *female* image and *male* image), to encode the perceived sex of the presented faces. There is a large literature on the relevance of sex/gender in face recognition (Mishra et al., 2019) and the dataset description mentions that 50% of the stimulus faces were female and 50% male. The sex of the study participants was recorded; it would also be possible to identify, record, and annotate the sex of the faces in the shared stimulus images. One could then ask questions as to whether sex of the imaged face influenced judgment response time or any other computed data feature.

Future analyses of the shared W-H dataset might address many possible questions: For example, many factors are known to generate effects on brain dynamics supporting face viewing and perception including left-right asymmetries, sex/gender effects, effects of portrait pose (straight-on versus angled), as well as repetition-related effects. The W-H experimental setup ensured that, to minimize head positional uncertainty, participants viewed the images straight on. However, the faces themselves were not all photographed in a full face pose. One might ask if there is a difference in the asymmetry classification of frontally photographed faces from the choice of pose. Did it take longer on average for participants to make symmetry judgments on non-frontal faces? Were there any effects of portrait pose angle on viewing-related brain source activities? Were there effects on source network dynamics? Time-on-task related factors might also be tested using the W-H dataset. Did the first run in a session have any statistically detectable differences from the last run? Was there a relationship of the handedness of the participant and the response time of the key presses corresponding to the assigned symmetry judgment? Was there any dependence of brain dynamics on the age of the portrait subject?

A thorough examination and annotation of potential experimental conditions *beyond* the basic design variables in the authors' conception of the experiment may also lead to a much more accurate assessment of bias, allowing systematic investigation of whether these biases influence computed data features of interest, and providing information enriching future mega-analysis across datasets by allowing it to consider effects of contrasts beyond those of primary concern to the original investigators.

4.6 Task specification

Guideline 5: All explicit or implicit participant tasks should be identified. *HED-3G defines a participant task as an organized participant activity performed during (or sometimes before or after) the experiment that may influence participant brain dynamics.* ***Explicit tasks*** *usually (though not always) determine and lead to actions that the participant performs (or inhibits) intentionally during the experiment – and* ***should always*** *be documented.* ***Implicit task challenges****, whether or not directly reflected in participant actions,* ***should also*** *be documented – particularly if they are part of the experimental design. Explicit pre- or post-session tasks*

external *to the recording session (often an aspect of experiments on learning or memory, for example) may also be considered for annotation, as in such experiment designs they may be intended to produce residual or priming effects in the session data.*

Explicit tasks. The W-H experiment has three instructed and thereby explicit tasks: face symmetry evaluation, gaze fixation, and blink-inhibition. The face symmetry evaluation task was the primary explicit task that the experimental participants were told to focus on. However, in the original data evaluation plan, this task was chosen solely to direct participant attention to each face and was irrelevant to the actual scientific goals of the experiment. Because this explicit task was the central activity the participant was instructed to perform, it therefore should be documented as an explicit task (even if, as here, it did not enter into the original data evaluation plan).

As is common with many MEEG experiments, the W-H experiment instructions also included two other explicit tasks: blink inhibition and gaze fixation. Participants were asked not to blink when a face was being shown and were also told to fixate their gaze on the cross when visible. Intentional fixation not only reduces the extent of natural eye movements but also may impose an additional mental load on participants. Instructed participant actions that may affect the recorded brain dynamics including, here, blink inhibition (Shultz et al., 2011) (Berman et al., 2012) and fixation (Stacchi et al., 2019), should always be considered explicit tasks for annotation. At a minimum, future analyses of the W-H dataset might test how successful participants were in inhibiting blinks during the specified period. Failures to inhibit might also be linked to variation in the recorded brain dynamics.

The separation of the two eye activity-related tasks into distinct tasks is necessary for the W-H dataset because the blink inhibition task was applied only while the face image is being displayed, while the gaze fixation task was instructed to be active during both the pre-stimulus interval and the face image presentation. Thus, these instructed intentions (affecting action) must be documented as separate tasks. While blink inhibition and gaze fixation could be annotated as experimental conditions in Table 2, activities performed intentionally by participants should be annotated as tasks, while elements that correspond to the setting and varying of experimental parameters should be annotated as experimental conditions or controls supporting interpretation of experiment control events and mega-analyses across datasets recorded under different conditions.

The W-H fMRI sessions also included a behavioral *face-memory test* conducted after the imaging session was completed. If the participant had been informed of this test before the session began (as often the case under informed consent), that foreknowledge might itself have had priming effects. However, since, in this case, participants did not have foreknowledge of the behavioral test, an experimental note to this effect should be included in the annotation of those data to inform further analysis. In the *face-memory test*, W-H asked participants to view face

images and to record whether they remembered seeing the face in the experiment sessions, by selecting one of three answers: (1) I did not ….; (2) I …; (3) I ….. These responses were not included in the original shared W-H dataset. To include them, BIDS standards expect that they be treated as a third, behavior-only, W-H experiment session. We plan to add these data to the HED-annotated MEEG data version on OpenNeuro.

Implicit tasks. The inclusion of repetition type as a design variable indicates that the experimenters were aware that detection of face novelty (or repetition) was very likely associated with brain dynamic effects in these data. The repetition type factor helps analysis assess the influence of this design factor in the data. The detection of *face novelty* can be considered to be an "implicit task", that is, an activity that the participants were not directly instructed to perform, but rather could be expected to perform (either intentionally or near-automatically) during the course of the experiment, or at very least, that could affect the recorded brain dynamics in some systematic manner. The repetition type design variable could also be associated with a *face recall* implicit task, recognizing that the brain dynamic correlates of both repeated-face recognition and new-face novelty detection are known to be associated with distinct brain activity patterns (Debener et al., 2005) (Murashko & Shmukler, 2019) (Courchesne et al., 1975).

The `face_type` design variable, indicating whether the image is of a famous face, an unfamiliar face, or shows a scrambled face, is also an obvious candidate here for implicit task designation. The mixed presentation of these three rather different sets of images can be expected to have posed one or more implicit task demands on most or all of the participants. Here, possible implicit tasks include *nonface recognition*, *known face recognition*, *unknown face appraisal*, and *known face identification*. Here the scrambled face (*nonface*) images were a (⅓) minority of the presented stimuli and differed markedly from the other face stimuli in visual presentation. MEEG responses to novel, outside-expected-category stimuli have distinct and long-known features.

Identifying the face image stimuli as either famous or anonymous here can support annotation of another implicit task. Again, both categories can be expected to have qualitatively different signatures in the recorded brain dynamics. In some cases, a famous face would be recognized as an individual known to the participant (implicit face identification), although in other cases the participant might vaguely recognize a presented face as famous without explicit retrieval of their personal identity. The (famous, anonymous) face annotations allow analysis tools to search for tasks involving specific cognitive processes – here, discriminating, identifying, or explicitly recalling information. By annotating such implicit tasks, shared datasets become amenable to future cross-dataset meta-analysis (of computed data features) and mega-analysis (of the raw data).

Clearly, there are potentially a large number of implicit tasks that could be annotated for analysis. The choice of how to identify and annotate implicit tasks depends on what the annotator

thinks may be profitable to look for in the data. Implicit task annotations direct downstream users of the data towards aspects of the experiment that are or may be associated with effects in the data. Very often, implicit tasks are associated with experimental control variables for experimental design or bias control. We anticipate that common practice here will develop gradually as researchers see the value added to their data by performing the annotation in a style compatible with other shared datasets involving different experiment and task designs.

The W-H fMRI sessions also included a behavioral *face-memory test* conducted after the imaging session was completed. If the participant had been informed of this test before the session began (as often the case under informed consent), that foreknowledge might itself have had priming effects. However, since, in this case, participants did not have foreknowledge of the behavioral test, an experimental note to this effect should be included in the annotation to inform further analyses.

4.7 Documenting temporal organization and architecture

Guideline 6: The temporal architecture of each recording should be annotated. *The internal* ***temporal architecture of each recording should be documented****, including timing of trials and rest periods between trials. If blocks of trials were used to vary or counterbalance some aspect of the experiment, event markers for the beginnings and ends of these blocks should also be included. Generally,* ***information that was fixed for the entire dataset should be gathered and annotated as a meta-event inserted at the time of the first data sample.***

Another part of the documentation important to proper MEEG data annotation is the specification of the experiment's temporal organization. Many MEEG datasets are organized into blocks of trials interspersed with rest periods. The W-H MEEG sessions were organized into 6 runs, each of 7.5 minutes duration and containing between 140 and 150 trials. The separate W-H fMRI sessions had more and shorter runs and also included 20-s fixation intervals to give repeated baseline data following each 50-s trial.

Within each run, the W-H MEEG data do not have an explicit block structure beyond the trial level, though other experiments may have temporal structure within runs imposed to counter-balance various experimental factors. A review of the W-H MEEG dataset metadata showed that between 3 and 6 minutes elapsed between MEEG session runs. Analysts assume that electrode caps or other sensors were not repositioned between runs *in the same session*. If this was *not* the case, the information should be clearly marked in the data, typically by separating it into separate *data sessions* in which channel locations do not (or are assumed to not) vary. Head movements with respect to the MEG dewar and its embedded sensors are a key concern in MEG studies, and movement files acquired at 1-second intervals are available for the W-H MEEG dataset.

Although the W-H experiment does not have a particularly complex temporal architecture, the authors do use the concept of an experimental trial, so a definition *(Definition/Trial, (Experimental-trial))* could be included to annotate the onset and offset of these trials if this would seem useful for planned analyses. Note however our cautions (Section 4.2) about annotating events only in relation to trial event groupings.

4.8 The event design process

Event design is usually an iterative process. Below are suggested steps to maximize the chances that the design leads to valuable and complete annotation:

1. ***Sketch a rough time-line similar to Fig. 1:*** having a good picture in mind of how the experiment unfolds is a very helpful starting point.
2. ***List the basic concepts of the experiment***: the sensory presentations, the participant motor and/or verbal responses, the tasks, the experimental design and bias control factors, and markers indicating the temporal structure of the dataset.
3. ***Write a concise but complete text description of each concept:*** a good starting point is to create a table of component names and descriptions (e.g., Table 2 and Table 11 below).
4. ***List the needed event marker types (e.g., Fig. 1).***
5. ***Assign a primary event category tag from the HED Event subtree to each marker (e.g., Fig. 2).***
6. ***Determine which additional columns should be in the events file.***
7. ***Verify that the experimental concepts (e.g., stimuli, responses, factors, levels, tasks) have appropriate HED definitions and can be associated with event markers or with event file columns.***
8. ***Iterate as necessary.***

The values in the `event_type` column indicate the types of event markers that should be placed in the data (cf. 1st column in Table 2). The **`HED`** column entries of Table 1 describe what occurs at each event marker in terms of important experimental concepts. Table 2 illustrates the process of creating definitions for important concepts. In performing event design, annotators should not try to fill in detailed HED tags at this stage, but should make sure that the relation of the events to the experiment structure is correctly expressed. Sensorimotor details can be easily added or modified later in the process, if using BIDS standards data by editing the `events.json` files. The first (**Name)** column of Table 11 illustrates which experimental concepts (in this case sensory presentations) are named in Step 2 of the procedure outlined above. The second column of Table 11 (**Concept description)** column gives an example text description (Step 3). The third column of Table 11 (**HED definition body)** column gives sample HED definitions for these concepts. Usually, this type of annotation occurs at the end of the design process after the event design has been determined.

5. Discussion and roadmap

Good event design and annotation are essential for ensuring the usability and longevity of both shared and stored MEEG data. Researchers need to think beyond the immediate problem to be analyzed and give thought as to how to share data in a manner that allows other researchers to rely on and benefit their research by its use. Many publishers are encouraging researchers to publish their data in a publication distinct from the primary published work. Separate publication not only increases the visibility of the work and provides authors with the opportunity to produce data with very high quality documentation. Separate publication provides an additional reward structure to researchers for sharing well-structured and well-annotated datasets and opens a potential venue for data authors to document and promote additional uses for their datasets beyond the primary publication.

Current standards and conventions for sharing neuroimaging data such as BIDS focus on file structure and inclusion of basic metadata, but have few requirements with respect to annotation of events. In fact, we know of no system other than HED that supports annotation of the nature of marked *events* in human neuroimaging time series data. Many of the BIDS validated MEEG datasets that we have evaluated on OpenNeuro have sparse or missing event annotations (Martínez-Cancino et al., 2020). For such BIDS datasets, adding a single `events.json` sidecar file, as illustrated here, or improving an existing one, may be all that is needed to turn an otherwise impoverished and unusable dataset into a richly informative one.

Annotators should begin by simply naming and describing sensory presentations, participant response actions, explicit tasks, and task conditions. Even without including very detailed HED tags in the definitions of these concepts, their presence in the annotation can allow future automated tools to produce detailed informative dataset summaries and structural information. Additional details can added to the JSON sidecar file at any times without modifying the rest of the dataset. We are working on HED tagging and use tutorials and plan to hold workshops in connection with major neuroimaging conferences to increase the visibility of the HED-3G approach to data sharing demonstrated here, and to involve the greater neuroscience community in advancing the use, capabilities and expressiveness of HED annotations.

Ideally, a thoughtful approach to event design as defined here should be initiated before the experiment begins. The reported event streams should be unwound so that each event is reported (*by-event*) in its own row in an `events.tsv` file rather than having some events being reported indirectly as offsets or response times relative to other reported events (Section 4.2). The latter (*by-trial*) approach can result in hopelessly convoluted event streams, particularly when additional data-feature or expert-annotation events are added *post hoc*. Such reporting makes analyses such

as regressing out the effects of overlapping temporal events nearly impossible without extensive manual re-coding specific to each dataset.

As laid out in this paper, the HED descriptor system provides a flexible framework for annotating events to any level of detail. Also, HED-3G now supports library schema, specialized HED vocabulary trees that can extend the vocabulary of the HED base schema for use by specific research user communities and applications. These library schemas can be used in conjunction with the HED base schema (viewable in an easy-to-use expanding format at *http://www.hedtags.org/display_hed*) to provide very powerful annotation capabilities. Currently, a SCORE library schema for labeling neurology clinical annotations used in neurology (Beniczky et al., 2017) is under development, and work is beginning on a MOVIE library schema for annotating experiments involving 4-D (animated) stimulus presentation. We are ready to assist other interested user community groups in developing library schemas to make available specialized subfield annotation vocabularies, available in HED, such as those needed to describe experiments involving music, linguistics, biomechanics, virtual reality, etc..

We also expect to make more progress on difficult remaining annotation issues including documenting spatial relationships, body movement frames, and task designs in HED. As mentioned, we also plan to work with experiment control program developers to investigate approaches for adding HED tags to experimental events and recorded participant actions during data acquisition. We look forward to documenting and demonstrating the value of the HED-3G context framework, only briefly discussed here (Section 2.4), for performing context-aware analysis of MEEG dynamics.

HED tools for validation and analysis support, some already implemented and others under development are being written in Python. A HED JavaScript validation tool has been incorporated into the official BIDS validator and is being continually improved. Online tools are available at https://hedtools.ucsd.edu. The CTagger annotation tool (https://github.com/hed-standard/CTagger) provides a simple-to-use interface that supports 'learning through doing' HED annotation. HED support tools are also being incorporated into EEGLAB and MATLAB, including tools to select and process data epochs based on searches through HED annotations. Further documentation is available on the HED website (https://www.hedtags.org). All HED code and issue forums are available on the HED organization GitHub website (https://github.com/hed-standard).

Finally, we should not ignore the suitability for HED-3G annotation to be applied equally well to events noted in other time series data including fMRI. Because of the much slower time constant of BOLD signals, fMRI research has recently to some extent moved away from using event-related designs. However, sensory presentations and participant actions, as well as in-data changes in experimental parameters and conditions in the many thousands of reported fMRI experiments are

equally well suited to HED annotation as are typically quite similar events in many MEEG experiments.

We believe that the time has now arrived for widespread recognition and acceptance of the need for a common framework for performing event annotation of neuroimaging time series data. The proposed HED-3G specification, now in beta version, will be officially released as version 8.0.0 soon. We welcome reader comments, suggestions, and participation going forward.

Acknowledgments:
We would like to express our deep appreciation to Daniel Wakeman and Richard Henson for sharing their rich and useful dataset with the human neuroscience research community, for their patience in helping us to understand the experimental details, and for their willingness to dig out additional information for us to include and discuss here. We would also like to acknowledge the seminal vision and contributions of Nima Bigdely-Shamlo, who initially conceived the HED system and demonstrated its practical use and importance for search and analysis across studies. We also thank Jonathan Touryan of the Army Research Laboratory and Tony Johnson of DCS Corporation for their work in supporting the development of HED for EEG data sharing. Ian Callanan and Alexander Jones are primary developers of HED-3G supporting infrastructure. This project received support from the Army Research Laboratory under Cooperative Agreement Number W911NF-10-2-0022 (KR) and from NIH projects R01 EB023297-03, R01 NS047293-l4, and R24 MH120037-01 (SM). The Swartz Center for Computational Neuroscience is supported in part by a generous continuing gift from The Swartz Foundation (Old Field, NY).

Author contributions:
Conceptualization: KR, SM, DT
Methodology: KR, SM, DT
Writing – original draft: KR, SM, DT, SA, AD
Writing – review & editing: KR, SM, DT, SA, AD
Data curation: DT, KR, AD
Software: KR, DT, AD, SA
Visualization: SM, KR

Conflicts of interest: The authors do not have any conflicts of interest.

Ethics: All data used in this study are publicly available.

Data availability: *https://openneuro.org/ds003645***.**

References

Beniczky, S., Aurlien, H., Brøgger, J. C., Hirsch, L. J., Schomer, D. L., Trinka, E., Pressler, R. M., Wennberg, R., Visser, G. H., Eisermann, M., Diehl, B., Lesser, R. P., Kaplan, P. W., Nguyen The Tich, S., Lee, J. W., Martins-da-Silva, A., Stefan, H., Neufeld, M., Rubboli, G., … Herman, S. T. (2017). Standardized computer-based organized reporting of EEG: SCORE – Second version. *Clinical Neurophysiology*, *128*(11), 2334–2346. https://doi.org/10.1016/j.clinph.2017.07.418

Berman, B. D., Horovitz, S. G., Morel, B., & Hallett, M. (2012). Neural correlates of blink suppression and the buildup of a natural bodily urge. *NeuroImage*, *59*(2), 1441–1450. https://doi.org/10.1016/j.neuroimage.2011.08.050

Bigdely-Shamlo, N., Cockfield, J., Makeig, S., Rognon, T., La Valle, C., Miyakoshi, M., & Robbins, K. A. (2016). Hierarchical Event Descriptors (HED): Semi-structured tagging for real-world events in large-scale EEG. *Frontiers in Neuroinformatics*, *10*. https://doi.org/10.3389/fninf.2016.00042

Bigdely-Shamlo, N., Kreutz-Delgado, K., Robbins, K., Miyakoshi, M., Westerfield, M., Bel-Bahar, T., Kothe, C., Hsi, J., & Makeig, S. (2013). Hierarchical Event Descriptor (HED) tags for analysis of event-related EEG studies. *2013 IEEE Global Conference on Signal and Information Processing*, 1–4. https://doi.org/10.1109/GlobalSIP.2013.6736796

Bigdely-Shamlo, N., Touryan, J., Ojeda, A., Kothe, C., Mullen, T., & Robbins, K. (2019). Automated EEG mega-analysis II: Cognitive aspects of event related features. *NeuroImage*, 116054. https://doi.org/10.1016/j.neuroimage.2019.116054

Boedhoe, P. S. W., Heymans, M. W., Schmaal, L., Abe, Y., Alonso, P., Ameis, S. H., Anticevic, A., Arnold, P. D., Batistuzzo, M. C., Benedetti, F., Beucke, J. C., Bollettini, I., Bose, A., Brem, S., Calvo, A., Calvo, R., Cheng, Y., Cho, K. I. K., Ciullo, V., … Twisk, J. W. R.

(2019). An empirical comparison of meta- and mega-analysis with data from the ENIGMA Obsessive-Compulsive Disorder Working Group. *Frontiers in Neuroinformatics*, *12*. https://doi.org/10.3389/fninf.2018.00102

Casson, A. J. (2019). Wearable EEG and beyond. *Biomedical Engineering Letters*, *9*(1), 53–71. https://doi.org/10.1007/s13534-018-00093-6

Costafreda, S. G. (2012). Meta-Analysis, Mega-Analysis, and Task Analysis in fMRI Research. *Philosophy, Psychiatry, & Psychology*, *18*(4), 275–277. https://doi.org/10.1353/ppp.2011.0049

Courchesne, E., Hillyard, S. A., & Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and Clinical Neurophysiology*, *39*(2), 131–143. https://doi.org/10.1016/0013-4694(75)90003-6

Debener, S., Makeig, S., Delorme, A., & Engel, A. K. (2005). What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis. *Cognitive Brain Research*, *22*(3), 309–321. https://doi.org/10.1016/j.cogbrainres.2004.09.006

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

Henson, R. N., Abdulrahman, H., Flandin, G., & Litvak, V. (2019). Multimodal Integration of M/EEG and f/MRI Data in SPM12. *Frontiers in Neuroscience*, *13*. https://doi.org/10.3389/fnins.2019.00300

Henson, R. N., Wakeman, D. G., Litvak, V., & Friston, K. J. (2011). A Parametric Empirical Bayesian Framework for the EEG/MEG Inverse Problem: Generative Models for Multi-Subject and Multi-Modal Integration. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00076

Holdgraf, C., Appelhoff, S., Bickel, S., Bouchard, K., D'Ambrosio, S., David, O., Devinsky, O., Dichter, B., Flinker, A., Foster, B. L., Gorgolewski, K. J., Groen, I., Groppe, D., Gunduz, A., Hamilton, L., Honey, C. J., Jas, M., Knight, R., Lachaux, J.-P., … Hermes, D. (2019). IEEG-BIDS, extending the Brain Imaging Data Structure specification to human intracranial electrophysiology. *Scientific Data*, *6*(1), 102. https://doi.org/10.1038/s41597-019-0105-7

Jas, M., Jones, S. R., & Hämäläinen, M. S. (2021). Whole-head OPM-MEG enables noninvasive assessment of functional connectivity. *Trends in Neurosciences*. https://doi.org/10.1016/j.tins.2021.04.006

Makeig, S., Gramann, K., Jung, T.-P., Sejnowski, T. J., & Poizner, H. (2009). Linking brain, mind and behavior. *International Journal of Psychophysiology*, *73*(2), 95–100. https://doi.org/10.1016/j.ijpsycho.2008.11.008

Martínez-Cancino, R., Delorme, A., Truong, D., Artoni, F., Kreutz-Delgado, K., Sivagnanam, S., Yoshimoto, K., Majumdar, A., & Makeig, S. (2020). The open EEGLAB portal interface:High-performance computing with EEGLAB. *NeuroImage*, 116778. https://doi.org/10.1016/j.neuroimage.2020.116778

Mishra, M. V., Likitlersuang, J., Wilmer, J. B., Cohan, S., Germine, L., & DeGutis, J. M. (2019). Gender differences in familiar face recognition and the influence of sociocultural gender inequality. *Scientific Reports*, *9*(1), 17884. https://doi.org/10.1038/s41598-019-54074-5

Murashko, A. A., & Shmukler, A. (2019). EEG correlates of face recognition in patients with schizophrenia spectrum disorders: A systematic review. *Clinical Neurophysiology*, *130*(6), 986–996. https://doi.org/10.1016/j.clinph.2019.03.027

Niso, G., Gorgolewski, K. J., Bock, E., Brooks, T. L., Flandin, G., Gramfort, A., Henson, R. N., Jas, M., Litvak, V., T. Moreau, J., Oostenveld, R., Schoffelen, J.-M., Tadel, F., Wexler, J., & Baillet, S. (2018). MEG-BIDS, the brain imaging data structure extended to magnetoencephalography. *Scientific Data*, *5*, 180110. https://doi.org/10.1038/sdata.2018.110

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

Robbins, K., Truong, D., Jones, A., Callanan, I., & Makeig, S. (2020). *Building FAIR functionality: Annotating event-related imaging data using Hierarchical Event Descriptors (HED)*. https://doi.org/10.31219/osf.io/5fg73

Shapiro, L. (2019). *Embodied Cognition*. Routledge. https://doi.org/10.4324/9781315180380

Shultz, S., Klin, A., & Jones, W. (2011). Inhibition of eye blinking reveals subjective perceptions of stimulus salience. *Proceedings of the National Academy of Sciences*, *108*(52), 21270–21275. https://doi.org/10.1073/pnas.1109304108

Squires, K., Petuchowski, S., Wickens, C., & Donchin, E. (1977). The effects of stimulus sequence on event related potentials: A comparison of visual and auditory sequences. *Perception & Psychophysics*, *22*(1), 31–40. https://doi.org/10.3758/BF03206077

Stacchi, L., Ramon, M., Lao, J., & Caldara, R. (2019). Neural representations of faces are tuned to eye movements. *Journal of Neuroscience*, *39*(21), 4113–4123. https://doi.org/10.1523/JNEUROSCI.2968-18.2019

Vitali, R. V., & Perkins, N. C. (2020). Determining anatomical frames via inertial motion capture: A survey of methods. *Journal of Biomechanics*, *106*, 109832. https://doi.org/10.1016/j.jbiomech.2020.109832

Wakeman, D. G., & Henson, R. N. (2015). A multi-subject, multi-modal human neuroimaging dataset. *Scientific Data*, *2*, 150001. https://doi.org/10.1038/sdata.2015.1

Supplementary materials

Supplementary Table 1: Complete JSON event file for Wakeman-Hansen dataset (short-form). The file has been re-spaced for readability. All of the definitions have been gathered into additional metadata dictionaries at the end of the file.

```
{
 "onset": {
   "Description": "Onset of the event in seconds relative to the start of the recording.",
   "Units": "s"
 },
 "duration": {
   "Description": "Duration of the event in seconds.",
   "Units": "s"
 },
 "event_type": {
   "LongName": "Event category",
   "Description": "The main category of the event.",
   "Levels": {
     "setup": "Mark start of experiment and document applicable metadata.",
     "show_cross": "Display only a white cross to mark start of trial and fixation.",
     "show_face": "Display a face to mark end of pre-stimulus and start of blink-inhibition.",
     "show_circle": "Display a white circle to mark end of the stimulus and blink inhibition.",
     "left_press": "Experimental participant presses a key with left index finger.",
     "right_press": "Experimental participant presses a key with right index finger.",
     "left_sym": "Pressing key with left index finger means a face with above average symmetry.",
     "right_sym": "Pressing key with right index finger means a face with above average symmetry."
       },
   "HED": {
     "setup": "Experiment-structure, (Def/Initialize-recording, Onset)",
     "show_cross": "Sensory-event, (Intended-effect, Cue), (Def/Cross-only, Onset),
                (Def/Fixation, Onset), (Def/Trial, Onset), (Def/Circle-only, Offset)",
     "show_face": "Sensory-event, Experimental-stimulus, (Def/Face-image, Onset),
               (Def/Blink-inhibition-task, Onset), (Def/Cross-only, Offset)",
     "show_circle": "Sensory-event, (Intended-effect, Cue), (Def/Circle-only, Onset),
               (Def/Face-image, Offset), (Def/Blink-inhibition-task, Offset),
               (Def/Fixation-task, Offset)",
     "left_press": "Agent-action, Participant-response, Def/Press-left-finger",
     "right_press": "Agent-action, Participant-response, Def/Press-right-finger",
     "left_sym": "Experiment-structure, (Def/Left-sym-cond, Onset)",
     "right_sym": "Experiment-structure, (Def/Right-sym-cond, Onset)"
       }
 },
 "face_type": {
   "Description": "Factor indicating type of face image being displayed.",
   "Levels": {
     "famous_face": "A face that should be recognized by the participants.",
     "unfamiliar_face": "A face that should not be recognized by the participants.",
     "scrambled_face": "A scrambled face image generated by taking face 2D FFT."
   },
   "HED": {
     "famous_face": "Def/Famous-face-cond",
     "unfamiliar_face": "Def/Unfamiliar-face-cond",
     "scrambled_face": "Def/Scrambled-face-cond"
   }
```

```
},
"repetition_type": {
  "Description": "Factor indicating whether this image has been already seen.",
  "Levels": {
    "first_show": "Factor level indicating the first display of this face.",
    "immediate_repeat": "Factor level indicating this face was the same as previous one.",
    "delayed_repeat": "Factor level indicating face was seen 5 to 15 trials ago."
  },
  "HED": {
    "first_show": "Def/First-show-cond",
    "Immediate_repeat": "Def/Immediate-repeat-cond",
    "delayed_repeat": "Def/Delayed-repeat-cond"
  }
},
"stim_file": {
  "Description": "Path of the stimulus file in the stimuli directory.",
  "HED": "(Image, Pathname/#)"
},
"hed_def_sensory": {
  "Description": "Metadata dictionary for gathering sensory definitions",
  "HED": {
    "Cross_only_def": "(Definition/Cross-only, (Visual, (Foreground-view, (White, Cross),
            (Center-of, Screen)), (Background-view, Black),
            Description/A white fixation cross on a black background in the center of the screen.))",
    "Face_image_def": "(Definition/Face-image, (Visual, (Foreground-view,
            ((Image, Face, Hair), Color/Grayscale), ((White, Cross), (Center-of, Screen))),
            (Background-view, Black), Description/A happy or neutral face in frontal
            or three-quarters frontal pose with long hair cropped presented as an
            achromatic foreground image on a black background with a white fixation cross superposed.))",
    "Circle_only_def": "(Definition/Circle-only, (Visual, (Foreground-view,
           ((White, Circle), (Center-of, Screen))), (Background-view, Black),
           Description/A white circle on a black background in the center of the screen.))"
  }
},
"hed_def_actions": {
  "Description": "Metadata dictionary for gathering participant action definitions",
  "HED": {
    "Press_left_finger_def": "(Definition/Press-left-finger, (Experimental-participant,
           (Index-finger, Left-side), (Press, Keyboard-key), Description/The participant presses
           a key with the left index finger to indicate a face symmetry judgment.))",
    "Press_right_finger_def": "(Definition/Press-right-finger, (Experimental-participant,
           (Index-finger, Right-side), (Press, Keyboard-key), Description/The participant presses
           a key with the right index finger to indicate a face symmetry evaluation.))"
  }
},
"hed_def_conds": {
  "Description": "Metadata dictionary for gathering experimental condition definitions",
  "HED": {
    "Famous_face_cond_def": "(Definition/Famous-face-cond, (Experimental-condition,
           Label/Face-type, (Image, (Face, Famous)),
           Description/A face that should be recognized by the participants))",
    "Unfamiliar_face_cond_def": "(Definition/Unfamiliar-face-cond, (Experimental-condition,
           Label/Face-type, (Image, (Face, Unfamiliar)),
           Description/A face that should not be recognized by the participants.))",
    "Scrambled_face_cond_def": "(Definition/Scrambled-face-cond, (Experimental-condition,
           Label/Face-type, (Image, (Face, Disordered)),
           Description/A scrambled face image generated by taking face 2D FFT.))",
    "First_show_cond_def": "(Definition/First-show-cond, (Experimental-condition,
```

```
            Label/Repetition-type, (First-item, Repetition-number/1),
            Description/Factor level indicating the first display of this face.))",
      "Immediate_repeat_cond_def": "(Definition/Immediate-repeat-cond, (Experimental-condition,
            Label/Repetition-type, (Next-item, Repetition-number/2),
            Description/Factor level indicating this face was the same as previous one.))",
      "Delayed_repeat_cond_def": "(Definition/Delayed-repeat-cond, (Experimental-condition,
            Label/Repetition-type, (Later-item, Repetition-number/2),
            Description/Factor level indicating face was seen 5 to 15 trials ago.))",
      "Left_sym_cond_def": "(Definition/Left-sym-cond, (Experimental-condition,
            Label/Key-assignment, ((Keyboard-key, Left-side), Symmetrical),
            ((Keyboard-key, Right-side), Asymmetrical),
            Description/Left finger key press means above average symmetry.))",
      "Right_sym_cond_def": "(Definition/Right-sym-cond, (Experimental-condition,
            Label/Key-assignment, ((Button, Right-side), (Behavioral-evidence, Symmetrical)),
            ((Keyboard-key, Left-side), (Behavioral-evidence, Asymmetrical)),
            Description/Right finger key press means above average symmetry.))"
    }
  },
  "hed_def_tasks": {
    "Description": "Metadata dictionary for gathering task definitions",
    "HED": {
      "Face_symmetry_evaluation_task_def": "(Definition/Face-symmetry-evaluation-task,
            (Task, Experimental-participant, (Look, Face), (Discriminate, (Face, Symmetric)),
            (Press, Keyboard-key), Description/Evaluate degree of
            image symmetry and respond with key press evaluation.))",
      "Blink_inhibition_task_def": "(Definition/Blink-inhibition-task,
            (Task, Experimental-participant, Inhibit-blinks,
            Description/Do not blink while the face image is displayed.))",
      "Fixation_task_def": "(Definition/Fixation-task, (Task, Experimental-participant,
            (Fixate, Cross), Description/Fixate on the cross at the screen center.))",
      "Face_novelty_detection_task_def": "(Definition/Face-novelty-detection-task,
            ((Task, Implicit), Experimental-participant, (Look, Face), (Detect, (Face, Novel)),
            Description/Recognize presentations of previously unseen face images-implicit task.))"
        }
  },
 "hed_def_structure": {
     "Description": "Metadata dictionary for gathering temporal setup definitions",
     "HED": {
         "Trial_def": "(Definition/Trial, (Description/Trial structure information.))"
     }
 },
 "hed_def_setup": {
     "Description": "Metadata dictionary for gathering setup definitions",
     "HED": {
         "Initialize_recording_def": "(Definition/Initialize-recording, (Description/Setup stuff.))"
     }
  }
 }
```

Supplementary Table 2: The assembled form of the HED annotation for the second event in Table 3 (as shown in Table 5) and in three different forms expanded by tools. Form 1 is the form that would normally appear in the `events.json` sidecar and be viewed. The tag strings have been re-spaced and partially bolded for readability.

Form 1: Short-form annotation of the sensory event corresponding to the first showing of famous face image `f032.bmp`. Definitions are unexpanded (as shown in Table 5).

Sensory-event, Experimental-stimulus, (Def/Face-image, Onset), (Def/Blink-inhibition-task, Onset), (Def/Cross-only, Offset), Def/Famous-face-cond, Def/First-show-cond, (Image, Pathname/f032.bmp)

Form 2: Long-form annotation of the sensory event corresponding to the first showing of famous face image `f032.bmp`. Definitions are unexpanded. Terms from form1 are shown in bold.

*Event/**Sensory-event**, Task-property/Task-event-role/**Experimental-stimulus**,*
*(Attribute/Informational/Def/**Face-image**,*
*Data-property/Spatiotemporal-property/Temporal-property/**Onset**),*
*(Attribute/Informational/Def/**Blink-inhibition-task**,*
*Data-property/Spatiotemporal-property/Temporal-property/**Onset**),*
*(Attribute/Informational/Def/**Cross-only**,*
*Data-property/Spatiotemporal-property/Temporal-property/**Offset**),*
*Attribute/Informational/Def/**Famous-face-cond**, Attribute/Informational/Def/**First-show-cond**,*
*(Item/Object/Man-made-object/Media/Visualization/**Image**,*
*Attribute/Informational/Metadata/**Pathname/f032.bmp**)*

Form 3: Short-form annotation of the sensory event corresponding to the first showing of famous face image `f032.bmp`. Definitions are expanded. The annotation has been manually indented to improve readability.

Sensory-event,
Experimental-stimulus,
*((Def-expand/**Face-image**,*
(Visual, (Foreground-view, ((Image, Face, Hair), Color/Grayscale),
((White, Cross), (Center-of, Screen))), (Background-view, Black),
*Description/A happy or neutral face in frontal or three-quarters frontal pose with long hair cropped presented as an achromatic foreground image on a black background with a white fixation cross superposed.)), **Onset**),*
*((Def-expand/**Blink-inhibition-task**,*
(Task, Experimental-participant, Inhibit-blinks,
*Description/Do not blink while the face image is displayed.)), **Onset**),*
*((Def-expand/**Cross-only**,*
(Visual, (Foreground-view, (White, Cross), (Center-of, Screen)), (Background-view, Black),
*Description/A white fixation cross on a black background in the center of the screen.)), **Offset**),*
*(Def-expand/**Famous-face-cond**, (Experimental-condition, Label/Face-type,*
(Image, (Face, Famous)), Description/A face that should be recognized by the participants)),
*(Def-expand/**First-show-cond**, (Experimental-condition,*
Label/Repetition-type,(First-item, Repetition-number/1),
Description/Factor level indicating the first display of this face.)),
*(**Image**, **Pathname/f032.bmp**)*

Form 4: Long-form annotation of the sensory event corresponding to the first showing of famous face image `f032.bmp`. Definitions are expanded. Top levels have been manually indented to improve readability.

*Event/**Sensory-event**,*
*Task-property/Task-event-role/**Experimental-stimulus**,*
*((Attribute/Informational/Def-expand/**Face-image**,*
(Attribute/Sensory/Visual, (Attribute/Sensory/Visual/View/Foreground-view,
((Item/Object/Man-made-object/Media/Visualization/Image,
Item/Biological-item/Anatomical-item/Body-part/Head/Face,
Item/Biological-item/Anatomical-item/Body-part/Head/Hair),
Attribute/Sensory/Visual/Color/Grayscale),
((Attribute/Sensory/Visual/Color/CSS-color/White-color/White,
Item/Object/Geometric-object/2D-shape/Cross),
(Attribute/Relational/Spatiotemporal-relation/Positional-relation/Center-of,
Item/Object/Man-made-object/Device/IO-device/Output-device/Display-device/Screen))),
(Attribute/Sensory/Visual/View/Background-view,
Attribute/Sensory/Visual/Color/CSS-color/Gray-color/Black),
Attribute/Informational/Description/A happy or neutral face in frontal or three-quarters frontal pose with long hair cropped presented as an achromatic foreground image on a black background with a white fixation cross superposed.)),
*Data-property/Spatiotemporal-property/Temporal-property/**Onset**),*
*((Attribute/Informational/Def-expand/**Blink-inhibition-task**,*
(Attribute/Organizational/Task, Agent-property/Agent-task-role/Experimental-participant,
Action/Move/Move-body-part/Move-head/Move-eyes/Inhibit-blinks,
Attribute/Informational/Description/Do not blink while the face image is displayed.)),
*Data-property/Spatiotemporal-property/Temporal-property/**Onset**),*
*((Attribute/Informational/Def-expand/**Cross-only**,*
(Attribute/Sensory/Visual,(Attribute/Sensory/Visual/View/Foreground-view,
(Attribute/Sensory/Visual/Color/CSS-color/White-color/White,
Item/Object/Geometric-object/2D-shape/Cross),
(Attribute/Relational/Spatiotemporal-relation/Positional-relation/Center-of,
Item/Object/Man-made-object/Device/IO-device/Output-device/Display-device/Screen)),
(Attribute/Sensory/Visual/View/Background-view,Attribute/Sensory/Visual/Color/CSS-color/Gray-color/Black),
Attribute/Informational/Description/A white fixation cross on a black background in the center of the screen.)),
*Data-property/Spatiotemporal-property/Temporal-property/**Offset**),*
*(Attribute/Informational/Def-expand/**Famous-face-cond**,*
(Attribute/Organizational/Experimental-condition, Attribute/Informational/Label/Face-type,
(Item/Object/Man-made-object/Media/Visualization/Image,
(Item/Biological-item/Anatomical-item/Body-part/Head/Face,
Attribute/Categorical/Categorical-judgment/Famous)),
Attribute/Informational/Description/A face that should be recognized by the participants)),
(Attribute/Informational/Def-expand/First-show-cond,
(Attribute/Organizational/Experimental-condition, Attribute/Informational/Label/Repetition-type,
(Attribute/Relational/Ordering-relation/First-item,
Data-property/Quantitative-property/Repetition-number/1),
Attribute/Informational/Description/Factor level indicating the first display of this face.)),
*(Item/Object/Man-made-object/Media/Visualization/**Image**,*
*Attribute/Informational/Metadata/**Pathname/f032.bmp**)*