12.1.2. Generic Data Reader

The Generic Data Reader (GDR) is used to read in EEG files of various formats for which no special reader exists (e.g. proprietary laboratory formats). The reader uses a header file which describes a single EEG. This file is an ASCII file with the extension ".vhdr". It will normally be given the same base name as the raw data EEG that is described in it. The header file is stored in the raw data folder of the workspace.

The format of the header file is based on the Windows INI format. It consists of sections of different names containing keynames and assigned values. Here is an extract of a header file:

```
Brain Vision Data Exchange Header File Version 1.0
; Data created from history path:
; P300b/Raw Data/Filters/Segmentation/BaselineCorrection/Average

[Common Infos]
DataFile=P300b_Average.dat
MarkerFile=P300b_Average.vmrk
DataFormat=ASCII
; Data orientation: VECTORIZED=ch1,pt1, ch1,pt2..., MULTIPLEXED=ch1,pt1, ch2,pt1 ...
DataOrientation=VECTORIZED
DataType=TIMEDOMAIN
NumberOfChannels=32
```

The first line identifies the header file and is mandatory.

A semicolon at the beginning of a line identifies a comment which is ignored by the reader. Blank lines are also ignored. A section is identified by a line with a term enclosed in square brackets. The header extract above, for example, contains the [Common Infos] section. A header file can contain any number of sections.

The next lines show some keynames in this section and the values that have been assigned to them. A keyname can only occur once in a section. Its meaning depends on the section in which it occurs. There must be no blank before or after the assignment operator (equal sign). Most predefined keynames have a predefined value which is used by the reader if a keyname is not found.

If you want to generate such a file, it is best to export any EEG with the aid of the Generic Data Export function. This creates a header which is compatible with the GDR. Set the parameters in such a way that the format of the exported file is as close as possible to that of the one to be imported. Now you can optimize the header to meet your specific requirements.

The various predefined sections with keynames, meaning and default values are listed below.

[Common Infos] This section contains general information on the EEG file.		
Keyname	**Meaning**	**Default value**
DataFile	Name of the EEG file. If the name does not contain a path, it is assumed that the EEG file is in the same folder as the header file. The placeholder *$b* can be used in the name. It is replaced by the base name of the header file when the file is read in. Example: The entry **DataFile=$b-EEG.dat** is interpreted for a header file named *Test.vhdr* as **DataFile=Test-EEG.dat**	None A value must be specified
MarkerFile	Optional marker file containing a list of markers assigned to the EEG. If no path is specified explicitly, the marker file is searched for in the folder in which the header file is located. The format of the marker file is explained further below. Here, too, it is possible to use the placeholder *$b*.	-
DataFormat	Possible values: **ASCII, BINARY**	ASCII
DataOrientation	Possible values: **VECTORIZED** First the file contains all data points for the first channel, followed by all data points for the second channel etc. **MULTIPLEXED** Here, all channels for every data point follow on from each other directly. The data structure is multiplexed.	MULTIPLEXED
DataType	Possible values: **TIMEDOMAIN** The data is in the time domain. **FREQUENCYDOMAIN** The data is in the frequency domain. **FREQUENCYDOMAIN_COMPLEX** The data exists as complex frequency values. Each real value is followed by an imaginary value. **TIMEFREQUENCYDOMAIN** The data exists in several layers, as in the case of the continuous wavelet transform, for	TIMEDOMAIN

	example. Each channel is represented by a vector of data at a point in time. **TIMEFREQUENCYDOMAIN_COMPLEX** This type corresponds to TIMEFRQUENCYDOMAIN except that here each value exists as a complex number.	
NumberOfChannels	Number of channels in the EEG file.	None A value must be specified
SamplingInterval	The sampling interval is specified in µs in the time domain and in hertz in the frequency domain.	None A value must be specified
Averaged	This indicates whether the data set which is to be read in has already been averaged. This is particularly relevant for the enabling and disabling of transforms on the Transformations menu. Possible values: **YES** Yes, the data set represents data that has been averaged. **NO** No, the data set represents data that has not been averaged.	NO
AveragedSegments	Number of segments included in the average. This value is only evaluated when "Averaged=YES" is set.	1
SegmentDataPoints	If the data is segmented evenly, then the number of data points per segment can be specified here.	0
SegmentationType	Like Averaged, this variable is relevant for the enabling and disabling of transforms on the Analyzer's Transformations menu. Possible values: **NOTSEGMENTED** The data set has not been segmented. **MARKERBASED** The data set has been segmented on the basis of one or more marker positions. All segments have the same length. **FIXTIME** Segmentation was based on fixed times. All segments have the same length.	NOTSEGMENTED
DataPoints	Number of data points in the EEG file.	0

	If no predefined value has been specified, the data is read to the end of the file. As far as binary data is concerned, the TrailerSize parameter can be set in the [Binary Infos] section as an alternative.	
Layers	Number of layers in a multilayer EEG of the type *TIMEFREQUENCYDOMAIN(_COMPLEX)*	1
LayerLowerLimit	Lower limit in multilayer data. In the case of the type *TIMEFREQUENCYDOMAIN(_COMPLEX)*, the unit is Hertz.	0
LayerUpperLimit	Upper limit in multilayer data	0
LayerFunction	Function that describes the intervals between the layers of multilayer data. Possible values: LINEAR Linear function LOGARITHMIC Logarithmic function	LINEAR

[ASCII Infos]
This section is only relevant if DataFormat in the [Common Infos] section was set to ASCII.

Keyname	**Meaning**	**Default value**
DecimalSymbol	Decimal symbol that is used in the EEG file. This symbol can be a point or comma. In the header file, the decimal symbol is always a point.	Point (.)
SkipLines	Number of header lines to be skipped.	0
SkipColumns	Number of columns to be skipped at the beginning of a line.	0

[Channel Infos]
This section lists the individual channels and their properties.

Keyname	**Meaning**	**Default value**
Ch<x> "x" stands for the channel number, i.e. the keyname for the first channel is Ch1, for the second channel Ch2 etc.	Individual properties for the channel are specified with commas between them: <Channel Name>,<Reference Channel Name>,<Resolution in µV> Example: `Ch1=Fp1,,1` Here, the first channel is named Fp1. The reference channel is assumed to be the common reference channel because no entry has been	<Channel Number>,,1,0 i.e. `Ch1=1,,1` for channel 1, for example

	made. Resolution is 1 µV. Resolution is the value by which the value of the data point is multiplied to convert it to µV.	

[Binary Infos] This section is only relevant if DataFormat in the [Common Infos] section was set to BINARY.		
Keyname	**Meaning**	**Default value**
BinaryFormat	Possible values: **IEEE_FLOAT_32** IEEE floating point format, single precision, 4 bytes per value. **INT_16** 16-bit signed integer **UINT_16** 16-bit unsigned integer	INT_16
ChannelOffset	Channel offset at which the data starts. This offset is only relevant to vectorized data. ChannelOffset and DataOffset can be used simultaneously.	0
DataOffset	Size of the offset in the file at which the actual data starts.	0
SegmentHeaderSize	If the data is segmented evenly, the size of the segment header can be input here in bytes.	0
TrailerSize	Size of the trailer of the EEG file in bytes. This parameter can be specified as an alternative to DataPoints in [Common Infos] in order to stop reading in the data before the end of the EEG file is reached.	0
UseBigEndianOrder	This only applies to integer formats. It specifies whether big endian order is used, i.e. whether the most significant byte in a number is stored first (Macintosh, Sun). Possible values: **YES** Yes, big endian order is in use. **NO** No, little-endian order is in use (corresponds to the Intel specification).	NO

Coordinates Coordinates are listed here.		
Keyname	**Meaning**	**Default value**
Ch<x> "x" stands for the	Coordinates of an individual channel in the form:	If the value is not listed here, the Analyzer uses

channel number, i.e. the keyname for the first channel is Ch1, for the second channel Ch2 etc.	<Radius>,<Theta>,<Phi> Example: `Ch1=1,-92,-72` The coordinate system of the Analyzer is described in Annex B.	the electrode name of the channel, searches for the coordinates in the 10/10 system and uses them. If the channel name is unknown, the coordinates are set internally to 0,0,0.

GDR-compatible marker file

The marker file is based on the same principle of sections and keynames as the header file.

It should be given the file name extension ".vmrk" and the same base name as the associated EEG file.

The first line identifies the marker file and is as follows:

```
Brain Vision Data Exchange Marker File Version 1.0
```

The various predefined sections with keynames, meaning and default values are listed below.

[Common Infos] This section contains general information on the marker file.		
Keyname	**Meaning**	**Default value**
DataFile	Name of the EEG file. If the name does not contain a path, it is assumed that the EEG file is in the same folder as the marker file. This information is not evaluated by the GDR.	-

[Marker Infos] The individual markers and their properties are listed in this section.		
Keyname	**Meaning**	**Default value**
Mk<x> Here, "x" stands for the marker number, i.e. the keyname for the first marker is Mk1, for the second marker Mk2, etc.	Individual properties for a marker are specified with commas between them: <Type>,<Description>,<Position>, <Points>,<Channel Number>,<Date> Example: `Mk1=Time 0,,26,1,0` Here, the first marker has the type "Time 0", no description, the position is at data point 26, the length is 1 data point, and the channel number is 0 which means that this marker relates to all channels.	-

	The date is optional. It is only evaluated if the marker type is *New Segment*. The date has the following format: 4 digits = Year 2 digits = Month 2 digits = Day 2 digits = Hour (24-hour system) 2 digits = Minute 2 digits = Second 6 digits = Microsecond Consequently time is broken down to the microsecond level. The following specification `19990311140312003012` means 11 March 1999, 14:03:12,003012	