

1 **Quantification of Phase-Amplitude Coupling in Neuronal Oscillations: Comparison of Phase-**
2 **Locking Value, Mean Vector Length, and Modulation Index**

3 Mareike J. Hülsemann, Dr. rer. nat.^{a,b,*}, Ewald Naumann, Dr. rer. nat.^a, Björn Rasch, Prof.^b

4 a University of Trier, Faculty I – Psychology, Department of General Psychology and Methodology,
5 Universitätsring 15, 54286 Trier, Germany

6 b University of Fribourg, Department of Psychology, Division of Cognitive Biopsychology and
7 Methods, Rue P.A. de Faucigny 2, 1701 Fribourg, Switzerland

8 ***Corresponding author:** Mareike Hülsemann, University of Fribourg, Department of Psychology,
9 Division of Biopsychology and Methods, Rue P.A. de Faucigny 2, 1701 Fribourg, Switzerland, Email:
10 mareike.huelsemann@unifr.ch

11 **E-mail addresses:** mareike.huelsemann@unifr.ch (M. J. Hülsemann), naumann@uni-trier.de (E.
12 Naumann), bjoern.rasch@unifr.ch (B. Rasch)

13 **Declaration of interest:** none.

14 **Abbreviations:** EEG: electroencephalography, MEG: magnetencephalography, PLV: phase-locking
15 value, MVL: mean vector length, MI: modulation index, FFT: Fast Fourier transform, SI:
16 synchronization index, ESC: envelope-to-signal correlation, BA: phase binning combined with
17 ANOVA, wPLF: weighted phase locking factor, GLM: general linear model, FIR: finite impulse
18 response, ANOVA: analysis of variance, S. D.: standard deviation, SEM: standard error, CFC: cross-
19 frequency coupling

20 **Funding:** This project has received funding from the Research Focus “Psychobiology of Stress” within
21 the research initiative of the state Rhineland-Palatinate by the Ministry of Science and from the
22 European Research Council (ERC) under the European Union's Horizon 2020 research and innovation
23 programme (grant agreement n° 677875). Both funding Institutions had no further role in the study
24 design, the collection, analysis, and interpretation of data, the writing of the manuscript, and the decision
25 to submit the paper for publication. The work was performed at Trier University, Faculty I – Psychology,
26 Department of General Psychology and Methodology.

27 **Highlights (3-5)**

- 28 • mean vector length is most sensitive for differentiating coupling strength
- 29 • modulation index is most robust to differences in data length, sampling rate and SNR
- 30 • phase-locking value and mean vector length cannot detect biphasic phase-amplitude coupling

31

32 **Abstract**

33 Phase-amplitude coupling is a promising construct to study cognitive processes in
34 electroencephalography (EEG) and magnetencephalography (MEG). Due to the novelty of the concept,
35 various measures are used in the literature to calculate phase-amplitude coupling. Here, performance of
36 the three most widely used phase-amplitude coupling measures – phase-locking value (PLV), mean
37 vector length (MVL), and modulation index (MI) – is thoroughly compared with the help of simulated
38 data. We combine advantages of previous reviews and use a realistic data simulation, examine
39 moderators and provide inferential statistics for the comparison of all three indices of phase-amplitude
40 coupling. Our analyses show that all three indices successfully differentiate coupling strength and
41 coupling width when monophasic coupling is present. While the mean vector length was most sensitive
42 to modulations in coupling strengths and width, biphasic coupling can solely be detected by the
43 modulation index. Coupling values of all three indices were influenced by moderators including data
44 length, signal-to-noise-ratio, and sampling rate when approaching Nyquist frequencies. The modulation
45 index was most robust against confounding influences of these moderators. Based on our analyses, we
46 recommend the modulation index for noisy and short data epochs with unknown forms of coupling. For
47 high quality and long data epochs with monophasic coupling and a high signal-to-noise ratio, the use of
48 the mean vector length is recommended. Ideally, both indices are reported simultaneously for one data
49 set.

50

51 **Keywords:** phase-amplitude coupling, cross-frequency coupling, phase-locking value, mean vector
52 length, modulation index, simulated EEG/MEG data

53

54 **1. Introduction**

55 Phase-amplitude coupling is a promising method to study cognitive processes (Jensen, 2006; Jensen and
56 Lisman, 1998; Lisman and Jensen, 2013; Vosskuhl et al., 2015). There is no convention yet of how to
57 calculate phase-amplitude coupling, but instead much heterogeneity of phase-amplitude calculation
58 methods used in the literature. Most of these are reasonable measures from a theoretical point of view.
59 To provide empirical evidence for choosing one of these measures over another, this work thoroughly
60 compares the performance of the three most widely used phase-amplitude coupling measures with the
61 help of simulated EEG data. The measures are the phase-locking value (PLV) by Mormann et al. (2005),
62 mean vector length (MVL) by Canolty et al. (2006), and modulation index (MI) by Tort et al. (2008).
63 From a historical viewpoint, the first amplitude modulations that have been detected are amplitude
64 fluctuations of specific frequency bands, becoming apparent in the fast Fourier transform (FFT) of
65 constituents of these signals (Burgess and Ali, 2002; Novak et al., 1992; Pfurtscheller, 1976). Because
66 the FFT approach can solely reveal that the amplitude of a higher frequency oscillates at a lower
67 frequency (characteristic of one signal), these amplitude modulations should not be misinterpreted to
68 account for true temporal coupling between the instantaneous phase of the lower frequency and the
69 amplitude envelope of the higher frequency (association between two signals and definition of phase-
70 amplitude coupling). Neither the lower frequency itself nor its instantaneous phase are extracted in this
71 approach.

72 Some of the most widely used phase-amplitude coupling measures today are the phase-locking value
73 [PLV] (Mormann et al., 2005), also called synchronization index [SI] by Cohen (2008), the mean vector
74 length [MVL] (Canolty et al., 2006), the modulation index [MI] (Tort et al., 2008), the envelope-to-
75 signal correlation [ESC] (Bruns and Eckhorn, 2004), the general linear model approach [GLM] (Kramer
76 and Eden, 2013; Penny et al., 2008), phase binning combined with analysis of variance (ANOVA) [BA]
77 (Lakatos et al., 2005), and the weighted phase locking factor [wPLF] (Maris et al., 2011). All of these
78 measures use the instantaneous phase and amplitude of band-pass filtered signals to calculate a measure
79 that represents coupling strength. However, conceptual ideas and mathematical principles differ
80 substantially between measures.

81 Several of these phase-amplitude coupling measures were compared with the help of simulated and real
82 data in four reviews. Tort et al. (2010) executed the most extensive comparison so far, including most
83 of the above listed measures and evaluating their performance pertaining to tolerance to noise, amplitude
84 independence (independence from the amplitude of the amplitude-providing frequency band),
85 sensitivity to multimodality, and sensitivity to modulation width. The modulation index, introduced by
86 the same group (Tort et al., 2008), is well-rated in all aspects while, amongst others, the phase-locking
87 value has poor ratings in all aspects. The mean vector length has good ratings in some aspects (e. g.
88 tolerance to noise), but weaknesses in others (e. g. amplitude dependence).

89 Penny et al. (2008) introduced the GLM approach and compared it to the phase-locking value, mean
90 vector length, and envelope-to-signal correlation in respect to noise level, coupling phase, data length,
91 sample rate, signal non-stationarity, and multimodality. They found that the methods discriminated
92 between data simulated with and without coupling to different extents, ranging from below chance level
93 to perfect discrimination. Performance of the measures differed under poor conditions (high noise, low
94 sampling rate, etc.), however, all measures performed equally well under good conditions (longer
95 epochs, less noise, etc.).

96 Kramer and Eden (2013) introduced a new GLM cross-frequency coupling measure. It proves to be
97 valid and performs equally well as the modulation index. The advantages of this method are that it can
98 be interpreted as percentage change in amplitude strength due to modulation. Additionally confidence
99 intervals are easily computed and the measure can detect biphasic coupling.

100 When Onslow et al. (2011), compared three phase-amplitude coupling measures (mean vector length,
101 modulation index, cross-frequency coherence), they found that “no one measure unfailingly out-
102 performed the others” (Onslow et al., 2011, p. 56). They concluded that each measure seems to be
103 particularly suited for specific data conditions. Mean vector length for example is suitable for noisy data,
104 exploratory analyses (analysing a broad frequency spectrum) and when the power of the amplitude
105 providing frequency band is low.

106 The above cited reviews do not point to a single optimal measure for calculating phase-amplitude
107 coupling. They rather show that most – but not all – of the used measures perform well and are equally
108 affected by various confounders. Despite the availability of manifold measures, 79 % of studies use the

109 phase-locking value adapted for phase-amplitude coupling, mean vector length, or modulation index
110 (Hülsemann, 2016). Why is this the case? The phase-locking value is derived from a long-used, phase-
111 phase coupling measure that is easily adapted for the purpose of phase-amplitude measurement. Its
112 familiarity in the scientific community might have promoted its application. Possibly the predominant
113 application of mean vector length is due to its mathematical directness. The modulation index is
114 conceptually intuitive.

115 The majority of reviews used very straightforward data simulation methods. Oftentimes, a sinusoidal
116 oscillation is constructed at a lower phase-providing frequency and at a higher amplitude-providing
117 frequency. Phase-amplitude coupling is introduced by multiplying both signals (cf. Onslow et al., 2011,
118 p. 52). Amplitude is then extracted from the so constructed signal and phase is extracted from the pure
119 sinusoidal oscillation of the lower frequency. White noise is added to both signals. There are two pitfalls
120 in this approach. Both sinusoidal signals reflect a plain prototype of phase-amplitude coupling, but in
121 real neuronal data, pure sinusoidal oscillation cannot be filtered; rather, frequency bands containing
122 different amounts of various frequencies are extracted. Second, white noise is added to the simulated
123 data, even though it is known that not white noise but Brownian noise is inherent to brain dynamics (He
124 et al., 2010; Miller et al., 2009).

125 Because none of the hitherto existing reviews simultaneously meet the requirements of realistic
126 simulation of EEG data, providing inferential statistics for comparison of the measures, investigating
127 moderators of phase-amplitude coupling, and including the three most widely used measures (phase-
128 locking value, mean vector length, and modulation index), a new comparison of these methods is
129 presented here. We aim to combine the best aspects of all previous reviews. EEG data is simulated rather
130 realistically according to the procedure described by Kramer and Eden (2013). The influence of several
131 moderators (multimodality, data length, sampling rate, noise level, modulation strength, and modulation
132 width) inspired by Tort et al. (2010) is investigated. Sensitivity and specificity of the phase-amplitude
133 coupling measures are checked according to the methods described in Onslow et al. (2011). For all these
134 comparisons, inferential statistics are provided.

135

136

137 2. Material and Methods

138 2.1. Simulation of EEG Data and Implementation of Phase-Amplitude Coupling

139 A characteristic of natural EEG data is the proportionality of its frequency spectrum to a power law $P(f)$
140 $\sim (1/f^\beta)$. Namely, the higher the frequency f , the weaker the amplitude $P(f)$. The exponent β defines the
141 strength of the amplitude decrease. White noise is defined by $\beta = 0$, pink noise by $\beta = 1$ and Brownian
142 (red) noise by $\beta = 2$. Different investigations have shown that the frequency spectrum of human brain
143 activity relates to Brownian (red) noise, with $2 < \beta < 3$ (He et al., 2010; Miller et al., 2009). Because of
144 this, Brownian noise was generated using MATLAB code provided by Zhivomirov (2013), in order to
145 simulate EEG data (Figure 1A).

146 Simulated data was then filtered at a low phase-providing frequency, from here on referred to as phase
147 time series, with a narrow bandwidth of 2 Hz. The same data was filtered at a high amplitude-providing
148 frequency, from here on referred to as amplitude time series, with a broad bandwidth. The exact
149 bandwidth of the amplitude time series should depend on the frequency of the phase time series (Berman
150 et al., 2012; Dvorak and Fenton, 2014). Because of this data was filtered, such that the sidebands of the
151 modulating frequency were always included (i. e. centre frequency of amplitude-providing frequency
152 band \pm upper boundary of phase-providing frequency band).

153 A zero-phase Hamming-windowed sinc finite impulse response (FIR) filter implemented in EEGLAB
154 (`pop_eegfiltnew.m`) was used. This function automatically chooses the optimal filter order and transition
155 band width for a precisely selectable filter bandwidth. Low frequency was set to 8 – 10 Hz and high
156 frequency to 50 – 70 Hz. Filtering can seriously distort raw data (Widmann et al., 2015), therefore only
157 continuous data was filtered and first and last samples, where edge artefacts can occur, were later on
158 discarded.

159 To introduce coupling, the procedure of Kramer and Eden (2013) was followed. A Hanning window
160 plus one (i.e. each data point of the Hanning window is added with one) was multiplied with the
161 amplitude time series. This multiplication of the Hanning window with the amplitude time series was
162 not done continuously, but centred at either the relative maxima (peaks) or the relative maxima and
163 minima (peaks and troughs) of the phase time series, in order to simulate monophasic and biphasic
164 coupling, respectively. Extremum times are chosen because they are easy to detect. They relate to phase

165 angles of 0° and $180^\circ/-180^\circ$. Phase-amplitude coupling measures would not change if the coupling were
166 to be introduced at another phase angle. The Hanning window itself is multiplied with the factor I to
167 graduate the intensity of phase-amplitude coupling. To double the amplitude of the time series at the
168 specified time $I = 1.0$ is chosen. $I = 0.0$ reflects no phase-amplitude coupling (i.e. not modulating the
169 amplitude time series). The length of the Hanning window was also modulated to simulate different
170 “widths” of phase-amplitude modulation. Parameters chosen for these moderators are specified below.
171 In a final step, additional noise was added to the phase and amplitude time series. Therefore, Brownian
172 noise of the same length was simulated, band-pass filtered at the same frequencies as the phase and
173 amplitude time series, and added to the original phase and modulated amplitude time series, respectively.
174 Frequency matched noise is disruptive to the modulated phase-amplitude coupling and therefore allows
175 to check for the robustness of the phase-amplitude coupling measures.

176 Subsequently, phase and amplitude were extracted from the correspondent time series via Hilbert
177 transform, using the Signal Processing Toolbox of MATLAB (The MathWorks, Inc). Then continuous
178 phase and amplitude time series were segmented. This was done to introduce data discontinuities, which
179 are present in real data as well. Filtering, Hilbert transform, and phase or amplitude extraction were
180 always conducted on continuous data, to prevent filtering or other artefacts in the later analysed data
181 epochs.

182 Data sets with a length of 42, 105, and 180 seconds were simulated. This amount of data is sufficient to
183 simulate 30 trials with a length of 400, 2500 and 5000 milliseconds plus additional 30 seconds to
184 introduce data discontinuities when segmenting the data. These parameters were chosen to mirror typical
185 properties of event-related EEG data: (1) at least 30 trials per unique condition for which phase-
186 amplitude coupling will be calculated (Luck, 2014), (2) trial length between 400 and 5000 milliseconds,
187 and (3) data discontinuities between trials. Sampling rate was set to 1000 Hz (Cohen, 2014). In addition,
188 simulated data was resampled to 500 Hz in order to investigate the influence of sampling rate. Noise
189 was scaled by the factor 0.9, 1.0, and 1.1 in order to simulate different signal-to-noise ratios. Scaling
190 factor 0.9, 1.0, and 1.1 correspond to a noise signal strength of 90 %, 100 %, and 110 % compared to
191 the data signal strength. Four modulation strengths were realised: $I = 0.0$ for no coupling and $I = 0.9$, I
192 $= 1.0$, and $I = 1.1$ for increasing coupling strength ($I = 1.0$ doubling the original amplitude strength).

193 These values lie within the range of former studies (e. g. Kramer and Eden, 2013). The length of the
194 Hanning Window ranged between 22.5 % and 27.5 % of one low frequency cycle to modulate different
195 “widths” of phase-amplitude modulation. This width is equivalent to about a quarter of one cycle and
196 therefore covers the peak (or trough) phases of that low frequency cycle. At these phases, amplitude of
197 the higher frequency was increased. All parameters were realised for mono- and biphasic coupling
198 (factor multimodality).

199

200 2.2. Measuring Phase-Amplitude Coupling

201 To calculate phase-amplitude coupling, first, raw data is band-pass filtered in the frequency bands of
202 interest. Second, the real-valued band-pass filtered signal is transformed into a complex-valued analytic
203 signal. Finally, phase or amplitude is extracted from the complex-valued analytic signal. All these steps
204 can essentially be implemented in MATLAB with four lines of code:

```
205 filtered_data = pop_eegfiltnew(raw_data, lower_frequency_bound, upper_frequency_bound);  
206 analytic_signal = hilbert(filtered_data);  
207 phase = phase(analytic_signal);  
208 amplitude = abs(analytic_signal);
```

209

210 2.2.1. Phase-Locking-Value by Mormann et al. (2005)

211 For the calculation of the phase-locking value, phase is extracted from the low frequency filtered analytic
212 signal and amplitude is extracted from the high frequency filtered analytic signal. The amplitude time
213 series is then again Hilbert transformed and phase is extracted from the “second” analytic signal. By
214 these steps, one obtains phase angles for both time series for each data (time) point. For each time point
215 the phase angle of the Hilbert transformed amplitude time series is subtracted from the phase angle of
216 the phase time series, obtaining phase angle differences.

217 These phase angle differences can be plotted in a polar plane as vectors of the length one with the angle,
218 representing the respective phase angle difference (Figure 1B, left panels). A constant phase lag between
219 both time series indicates phase-amplitude coupling. A constant phase lag leads to vectors in the polar
220 plane with a similar direction. Then all vectors are averaged: if they have a constant phase lag, they
221 point into the same direction leading to a rather long mean vector. If there is a variable phase lag, the

222 vectors are scattered around the polar plane, leading to a rather short mean vector. The length of the
223 mean vector indicates the amount of phase-amplitude coupling (coupling strength). The direction of the
224 vector represents the mean phase lag present between the two time series and the preferred coupling
225 phase can be inferred from the phase lag. The phase-locking value is calculated by the following
226 formula:

$$227 \quad PLV = \left| \frac{\sum_{t=1}^n e^{i(\theta_{lt} - \theta_{ut})}}{n} \right| \quad \text{(Equation 1)}$$

228 where n is the total number of data points, t is a data point, θ_{lt} is the phase angle of the lower frequency
229 band at time point t and θ_{ut} is the phase angle of the Hilbert transformed upper frequency band amplitude
230 time series.

231 The logic for this measure is that if and only if the amplitude of the high frequency time series oscillates
232 at the lower frequency (indicator for phase-amplitude coupling) extracting instantaneous phase
233 information from this signal will return valid phase angles that may have a constant phase lag to the
234 instantaneous phase information of the low frequency band. If the amplitude of the high frequency time
235 series does not oscillate at the lower frequency band (indicator for lack of phase-amplitude coupling),
236 distorted phase information will be extracted from Hilbert transformed amplitude time series that will
237 have an inconsistent phase lag to the instantaneous phase of the lower frequency signal.

238 One should be aware, that meaningful phase information can only be extracted from narrow band
239 oscillations. The Hilbert transformed amplitude time series does not necessarily need to be such a narrow
240 band oscillation.

241

242 2.2.2. Mean Vector Length by Canolty et al. (2006)

243 The phase-amplitude coupling measure mean vector length (MVL) introduced by Canolty et al. (2006)
244 utilizes phase angle and magnitude of each complex number (i. e. each data point) of the analytic signal
245 in a quite direct way to estimate the degree of coupling. Each complex value of the analytic time series
246 is a vector in the polar plane. Phase-amplitude coupling is present, when the magnitude M of a fraction
247 of all vectors is especially high at a specific phase or at a narrow range of phases (Figure 1B, middle
248 panels). Averaging all vectors creates a mean vector with a specific phase and length (red vector in
249 Figure 1B). The length of this vector represents the amount of phase-amplitude coupling. The direction

250 represents the mean phase where amplitude is strongest. When no coupling is present, all vectors cancel
251 each other out and the mean vector will be short. Then its direction does not represent any meaningful
252 phase. The mean vector length is calculated by the following formula:

$$253 \quad \text{MVL} = \left| \frac{\sum_{t=1}^n a_t e^{i\theta_t}}{n} \right| \quad (\text{Equation 2})$$

254 where n is the total number of data points, t is a data point, a_t is the amplitude at time point t and θ_t is
255 the phase angle at time point t . This value cannot become negative because it represents the length of
256 the mean vector. The length of a vector cannot be negative.

257 Three caveats come along with this measure: (1) the value is dependent on the general absolute
258 amplitude of the amplitude providing frequency (independent of outliers), (2) amplitude outliers can
259 strongly influence the mean vector length, and (3) phase angles are often not uniformly distributed
260 (Cohen, 2014). All caveats are simultaneously counteracted by nonparametric permutation testing (see
261 section 2.2.4). One of the reviews cited in the introduction (Tort et al., 2010) finds faults with the mean
262 vector length being amplitude dependent. However, this is only true for the raw, but not for the permuted
263 mean vector length.

264 In the interest of completeness, it should be mentioned that Özkurt and Schnitzler (2011) proposed a
265 direct mean vector length which is amplitude-normalized and ranges between 0 and 1. When applying
266 permutation testing to both mean vector length and direct mean vector length return essentially the same
267 values. That is, when applied along with permutation testing, both measures are exchangeable. Without
268 permutation testing, the usage of the direct mean vector length is recommended because it takes care of
269 the possible amplitude differences in raw data.

270

271 2.2.3. Modulation Index by Tort et al. (2008)

272 Tort et al. (2008) suggests a very different way of computing phase-amplitude coupling, which anyways
273 is based on the same parameters of the analytic signal, amplitude magnitude and phase angle. For
274 calculating the modulation index (MI) according to Tort et al. (2008), all possible phases from -180° to
275 180° are first binned into a freely chosen amount of bins. Tort et al. (2008) established to use 18 bins of
276 20° each, which many authors follow. The amount of bins can influence the results, as will be explained

277 below. The average amplitude of the amplitude-providing frequency in each phase bin of the phase-
278 providing frequency is computed and normalized by the following formula:

$$279 \quad p(j) = \frac{\bar{a}}{\sum_{k=1}^N \bar{a}_k} \quad (\text{Equation 3})$$

280 where \bar{a} is the average amplitude of one bin, k is the running index for the bins, and N is the total amount
281 of bins; p is a vector of N values. With the help of these calculations, one obtains the data for the phase-
282 amplitude plot, which depicts the actual phase-amplitude coupling graphically (Figure 1B, right panels).
283 Subsequently Shannon entropy is computed; a measure that represents the inherent amount of
284 information of a variable. If Shannon entropy is not maximal, there is redundancy and predictability in
285 the variable. Shannon entropy is maximal, if the amplitude in each phase bin is equal (uniform
286 distribution, Figure 1B, right upper panel). Shannon entropy is computed by the following formula:

$$287 \quad H(p) = - \sum_{j=1}^N p(j) \log p(j) \quad (\text{Equation 4})$$

288 where p is the vector of normalized averaged amplitudes per phase bin and N is the total amount of bins.
289 It does not matter which logarithm base is used if permutation testing is applied later on (Cohen, 2014).
290 Like in Tort et al. (2008) the natural logarithm is used here. Shannon entropy is dependent on the amount
291 of bins used and this is why the modulation index is likewise dependent on the number of bins. The
292 higher the amount of bins, the larger Shannon entropy can become. Complying with the original author
293 and most other studies, 18 bins have been employed here.

294 Phase-amplitude coupling is defined by a distribution that significantly deviates from the uniform
295 distribution. Kullback-Leibler distance, a measure for the disparity of two distributions is calculated by
296 the following formula:

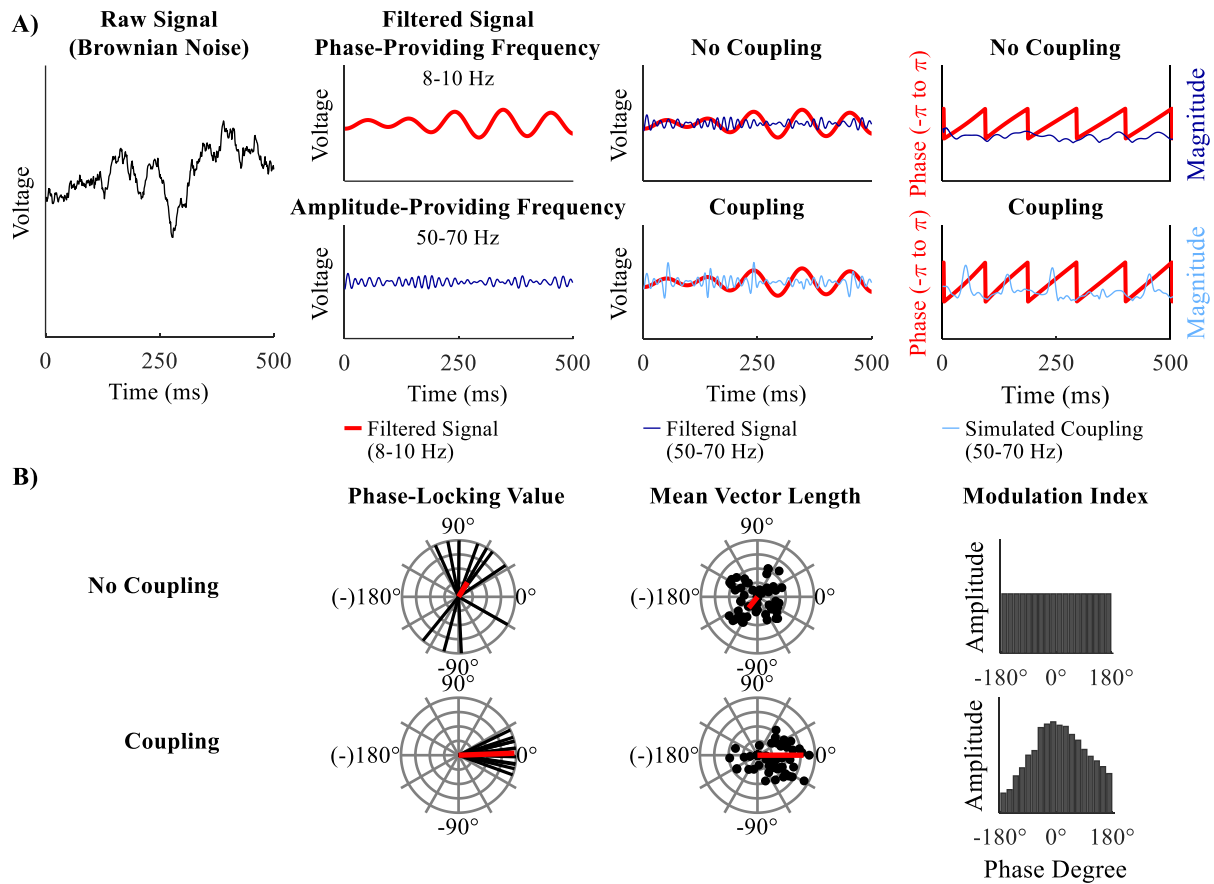
$$297 \quad KL(U, X) = \log N - H(p) \quad (\text{Equation 5})$$

298 where U is the uniform distribution, X is the distribution of the data, N is the total amount of bins, and
299 $H(p)$ is the Shannon entropy according to equation 4. The uniform distribution is represented by $\log(N)$.
300 The final raw modulation index is calculated by the following formula:

$$301 \quad MI = \frac{KL(U, X)}{\log N} \quad (\text{Equation 6})$$

302 where $KL(U, X)$ is the Kullback-Leibler distance according to equation 5 and N is the total amount of
303 bins.

304



305

306

[2 column fitting image]

307 **Figure 1 – Simulation of the EEG signal and calculation of phase-amplitude coupling:** A) (from
 308 left to right) Brownian noise is generated. This signal is band pass filtered to extract the slow phase-
 309 providing frequency (here 8-10 Hz, red line) and the fast amplitude-providing frequency (here 50-70
 310 Hz, dark blue line). To simulate coupling (light blue line) the amplitude-providing band pass filtered
 311 signal is multiplied with a Hanning window plus one (not depicted here), which results in stronger
 312 amplitude at the peaks of the phase-providing frequency (lower middle right panel). Before extracting
 313 phase and amplitude (most right panels) band pass filtered noise (same frequencies) is added to the
 314 filtered data (not depicted here). The simulated coupling (light blue line) amplitude is most pronounced
 315 for phases at 0°. This is not the case for the original signal (dark blue line). B) Idealized depiction phase-
 316 locking value (left panels), mean vector length (middle panels), and modulation index (right panels)
 317 for a uniform distribution (upper panels) and phase-amplitude coupling (lower panels). Phase-Locking
 318 Value: Each black line represents the phase lag between two signals at one time point. The red vector is

319 the mean of all black vectors. The upper panel shows inconsistent, widespread phase lags. The
320 widespread phase lags lead to a relatively short mean vector (red line). The left panel shows an example
321 of a relative constant phase lag around 0° . A relative constant phase lag leads to a relatively long mean
322 vector (lower panel). Mean Vector Length: Each black dot represents one data point of the analytical
323 signal. In case of coupling, a portion of the dots (or vectors) are especially long (reflecting strong
324 amplitudes) at a specific narrow range of phase angles (here 0° in the lower panel). The red vector is the
325 mean of all black vectors. It reflects coupling strength (short for no coupling – long for coupling). In
326 case of phase-amplitude coupling it is indicating the preferred phase. Modulation Index: All possible
327 phases are binned into 18 bins of 20° from -180° to 180° . Each bar reflects the mean amplitude of the
328 amplitude-providing signal for the specified phase of the phase-providing frequency. This phase-
329 amplitude plot is quantified with Shannon entropy. Shannon entropy is maximal for uniform
330 distributions (upper panel). The Kullback-Leibler distance measures how much a given distribution (for
331 example the one in the lower panel) deviates from the uniform distribution (depicted in the upper panel).
332 The more phase-amplitude coupling there is in the data, the more the given phase-amplitude plot
333 deviates from the uniform distribution and the higher the modulation index becomes.

334

335 2.2.4. Permutation Testing

336 All methods are subjected to permutation testing in order to quantify the meaningfulness of the derived
337 value (Cohen, 2014). For permutation testing, the observed coupling value is compared to a distribution
338 of shuffled coupling values. Shuffled coupling values are constructed by calculating the coupling value
339 between the original phase time series and a permuted amplitude time series (or vice versa). The
340 permuted amplitude time series is constructed by cutting the amplitude time series at a random time
341 point and reversing the order of both parts. Generating surrogate data this way is most conservative,
342 because it leaves all characteristics of the EEG data intact, except the studied one, namely the temporal
343 relationship between phase angle and amplitude magnitude. Shuffling is usually repeated 200 to 1000
344 times (here we used 1000). The observed coupling value is standardized to the distribution of the
345 shuffled coupling values according to the following formula:

$$346 \quad CV_z = \frac{CV_{\text{observed}} - \mu_{CV_{\text{shuffled}}}}{\sigma_{CV_{\text{shuffled}}}} \quad (\text{Equation 7})$$

347 where CV denotes coupling value, μ denotes the mean and σ denotes the standard deviation (S. D.).
348 Only when the observed phase-locking value is larger than 95 % of shuffled values (which are expected
349 to be uncorrelated), it is defined as significant.

350

351 2.3. Statistical Analyses

352 All statistical analyses were conducted with IBM Statistics for Windows Version 23 (SPSS, Inc., IBM
353 company), except otherwise specified. Significance level were set to $p < .05$. Violations of sphericity
354 were, whenever appropriate corrected by Greenhouse-Geisser ϵ (Geisser and Greenhouse, 1958).
355 Further analyses of significant results were conducted post hoc with Dunn's multiple comparison
356 procedure (Dunn, 1961) or post hoc t-tests. Effect size measure ω^2 is reported for significant results
357 (Hays, 1973). It is an estimator for the population effect Ω^2 , which specifies the systematic portion of
358 variance in relation to the overall variance (Rasch et al., 2006).

359

360 2.3.1. Specificity of phase-amplitude coupling measures

361 In a first step 10 000 data sets without coupling were simulated by setting the modulation strength to I
362 = 0. Simulations were carried out for the frequency pair 8 – 10 Hz for phase time series and 50 – 70 Hz
363 for amplitude time series. Phase-amplitude coupling values were generally compared in a 3 x 3 x 2 x 3
364 analysis of variance (ANOVA) with the repeated measurement factors method (phase-locking value,
365 mean vector length, modulation index), data length (400 ms, 2500 ms, 5000 ms), sampling rate (500 Hz,
366 1000 Hz), and noise level (90 %, 100 %, 110 %).

367 As described above, nonparametric permutation testing was performed. Raw phase-amplitude coupling
368 measures were z-standardized to the shuffled phase-amplitude coupling distribution. Normal z-values
369 directly imply p-values; a value of 1.64 corresponds to a p-value of 5 %. The phase-amplitude coupling
370 value distribution which is expected under the null-hypothesis does not have to match the standardised
371 normal distribution. Therefore, significance was not inferred from the standardised normal distribution,
372 but instead by that phase-amplitude coupling value, at which 5 % of simulated data (with no coupling)

373 was classified as false positive. Shuffling for permutation testing was done within trials. Coupling
374 measures were then calculated on concatenated trials.

375 Specificity of measures was analysed by counting false positives (significant coupling, even though it
376 was not engineered into the simulated data) depending on (1) method, (2) data length, (3) sampling rate,
377 and (4) noise level. To be able to conduct an ANOVA, the 10 000 simulations were divided into 100
378 subsamples of 100 simulations each. For each subsample false positives were counted. Each subsample
379 was treated as a case in the subsequent 3 x 2 x 3 x 2 ANOVA with the repeated measurement factors
380 method (phase-locking value, mean vector length, modulation index), data length (400 ms, 2500 ms,
381 5000 ms), sampling rate (500 Hz, 1000 Hz), and noise level (90 %, 100 %, 110 %) and the dependent
382 variable false positives.

383

384 *2.3.2. Sensitivity of phase-amplitude coupling measures as a function of moderating variables*

385 Performance of phase-amplitude coupling measures were quantified by simulating 100 independent data
386 sets and modifying the parameters (1) modulation strength, and (2) modulation width, (3) multimodality,
387 (4) data length, (5) sampling rate, and (6) noise level within each dataset. Six 2-way ANOVAs were
388 calculated. Each ANOVA included the repeated measurement factor method and was individually
389 combined with the repeated measurement factors modulation strength (90 %, 100 %, 110 %),
390 modulation width (22.5 %, 25.0 %, 27.5 % of one low frequency cycle), multimodality (monophasic,
391 biphasic), data length (400 ms, 2500 ms, 5000 ms), sampling rate (500 Hz, 1000 Hz), and noise level
392 (90 %, 100 %, 110 % compared to signal strength)

393

394

395 **3. Results and Discussion**

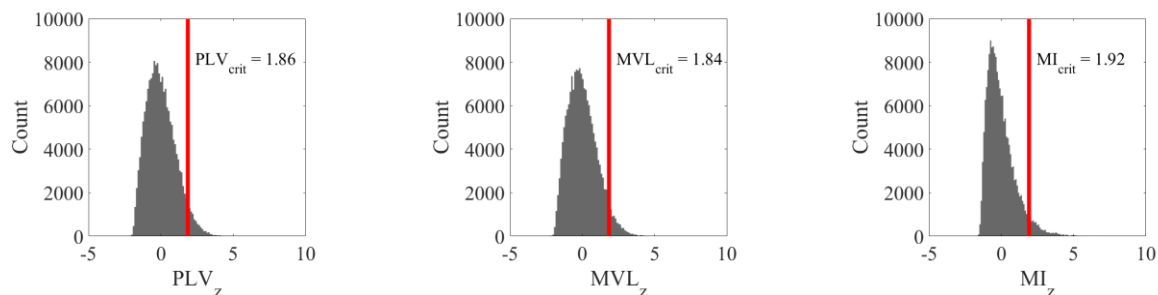
396 *3.1. Specificity of Phase-Amplitude Coupling Measures*

397 Phase-amplitude coupling values did not differ depending on method, data length, sampling rate, or
398 noise level. Because of the high number of simulations (n = 10 000), some main effects and interactions

399 became significant. However, all effect sizes were below $\omega^2 < .01$, therefore these differences are
400 negligible.

401 Figure 2 shows the phase-amplitude coupling value distribution for the phase-locking value, the mean
402 vector length, and the modulation index. When setting the critical z-value for the phase-locking value
403 at 1.86, for the mean vector length at 1.84, and for the modulation index at 1.92 five percent of the
404 simulated data were classified as containing coupling (false positive). Thus, these values were defined
405 as critical z-values. This implies that the mean vector length is most specific, directly followed by the
406 phase-locking value. The modulation index is least specific compared to the two other methods.

407



408

409

[2 column fitting image]

410 **Figure 2 – Probability distribution of coupling values under the null hypothesis:** Phase-amplitude
411 coupling value distribution under the null hypothesis (i. e. no coupling present in the data) of phase-
412 locking value (left panel), mean vector length (centre panel), and modulation index (right panel). These
413 distributions allow defining the significance threshold. The red line marks the critical phase-amplitude
414 coupling z-value (relative cut off of 5 %). Choosing an absolute cut off instead would lead to smallest
415 amount of false positives for mean vector length, followed by the phase-locking value. The modulation
416 index would detect the most false positives.

417

418 The amount of false positives did differ depending on data length ($F(2,198) = 27.19, p < .01, \omega^2 = .15,$
419 $Dunn_{crit} = .26$). There were significantly more false positives during short epochs (400 ms; mean \pm S.E:
420 $5.43 \pm .07$) compared to medium (2500 ms; mean \pm S.E: $4.70 \pm .06$) and long epochs (5000 ms; mean \pm
421 S.E: $4.78 \pm .09$). Medium and long epochs did not differ in their false positive rates.

422 The main effect was qualified by a method by data length interaction ($F(4,396) = 36.34, p < .01, \omega^2 =$
423 $.14, \text{Dunn}_{\text{crit}} = .23$). This revealed that the above-described pattern was driven by the phase-locking value
424 and mean vector length. There were no differences in false positive rate within the modulation index.
425 Furthermore, in short epochs there were significantly more false positive in phase-locking value and
426 mean vector length compared to the modulation index. In medium and long epochs there were
427 significantly less false positive in phase-locking value and mean vector length compared to the
428 modulation index.

429 Independently of the method, the main effect was further qualified by a sampling rate by data length
430 interaction ($F(2,198) = 36.14, p < .01, \omega^2 = .10, \text{Dunn}_{\text{crit}} = .32$). The above-described pattern of the main
431 effect was only evident for a sampling rate of 500 Hz. Furthermore, the interaction revealed that there
432 were more false positives for 500 Hz compared to 1000 Hz sampling rate during short epochs, no
433 difference in false positives between sampling rates in medium epochs, and less false positives for 500
434 Hz compared to 1000 Hz sampling rate during long epochs.

435

436 *3.2. Sensitivity of Phase-Amplitude Coupling Measures as a Function of Moderating Variables*

437 *3.2.1. Effect of method on phase-amplitude coupling measures*

438 Phase-locking value ($1.66 \pm .06$) and mean vector length ($2.08 \pm .07$) differed from the modulation index
439 ($11.97 \pm .75$) in their absolute magnitude independently of any other factor (main effect method:
440 $F(2,198) = 215.22, p < .01, \omega^2 = .59, \text{Dunn}_{\text{crit}} = 1.34$). Phase-locking value and mean vector length did
441 not differ from each other.

442

443 *3.2.2. Effect of modulation strength on phase-amplitude coupling measures*

444 Coupling values of all methods increased with increasing modulation strength ($F(2,198) = 189.05, p <$
445 $.01, \omega^2 = .56$). The interaction method by modulation strength became significant ($F(4,396) = 151.54, p$
446 $< .01, \omega^2 = .40$; Figure 3A). Post hoc t-tests showed that all factor levels within a method differed
447 significantly from each other (all p 's $< .01$). The effect of modulation strength was most pronounced for
448 the mean vector length ($.47 < \omega^2 < .76$), followed by the modulation index ($.49 < \omega^2 < .69$). The phase-
449 locking value was least sensitive to modulation strength ($.37 < \omega^2 < .72$).

450 The stronger the coupling, the larger phase-locking value, mean vector length, and modulation index
451 are. As Tort et al. (2010) has shown, this behaviour is not inherent to all phase-amplitude coupling
452 measures. Since researchers do not only want to prove the existence of phase-amplitude coupling, but
453 also differentiate its strength, a measure that can do this is indispensable. Of all three methods, mean
454 vector length differentiates best between the different factor levels of modulation strength.

455

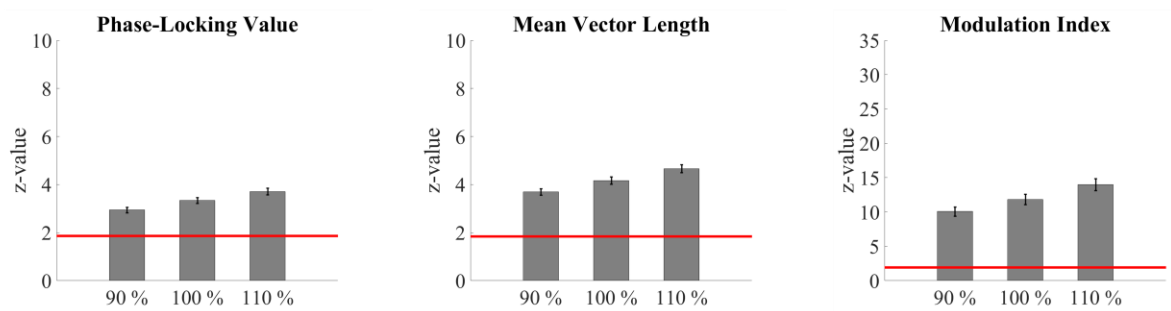
456 3.2.3. Effect of modulation width on phase-amplitude coupling measures

457 Coupling values of all methods increased with increasing modulation width ($F(2,198) = 110.11, p < .01,$
458 $\omega^2 = .42$). The interaction method by modulation width became significant ($F(4,396) = 70.18, p < .01,$
459 $\omega^2 = .24$; Figure 3B). Post hoc t-tests showed that all factor levels within a method differed significantly
460 from each other (all p 's $< .01$). The effect of modulation width was most pronounced for the phase-
461 locking value ($.14 < \omega^2 < .72$) and the mean vector length ($.14 < \omega^2 < .71$). The modulation index was
462 least sensitive to modulation width ($.15 < \omega^2 < .52$).

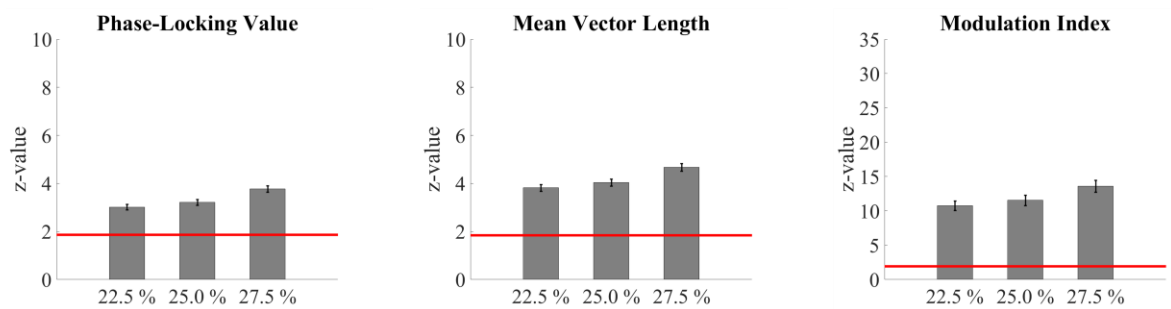
463 The broader the coupling width, the larger phase-locking value, mean vector length, and modulation
464 index are. Of all three methods, phase-locking value and mean vector length differentiate best between
465 the different factor levels of modulation width.

466

A) Modulation Strength Effect



B) Modulation Width Effect



467

468

[2 column fitting image]

469 **Figure 3 – Sensitivity for modulation strength and width:** Mean (\pm SEM) phase-amplitude coupling

470 values for each method for the A) modulation strength effect and B) modulation width effect. Coupling

471 values of all methods increased with increasing modulation strength. However, mean vector length

472 differentiates best between the different factor levels of modulation strength. Also, coupling values of

473 all methods increased with increasing modulation width. Here, phase-locking value and mean vector

474 length differentiate best between the different factor levels of modulation width. The red line marks the

475 significance level. All values above this line represent significant phase-amplitude coupling. For each

476 effect, all factor levels within a method are significantly different from each other according to post-hoc

477 t-tests. Only monophasic coupling values are depicted for the phase-locking value and the mean vector

478 length.

479

480 3.2.4. Effect of multimodality on phase-amplitude coupling measures

481 Monophasic coupling ($7.16 \pm .36$) led to overall stronger coupling measures than biphasic coupling

482 ($3.31 \pm .24$; $F(1,99) = 813.94$, $p < .01$, $\omega^2 = .80$). Biphasic coupling could not be detected by the phase-

483 locking value ($3.33 \pm .12$ vs. $-.01 \pm .01$; $t(99) = 27.26$, $p < .01$, $\omega^2 = .88$) and mean vector length (4.17

484 $\pm .15$ vs. $-01. \pm .01$; $t(99) = 27,85$, $p < .01$, $\omega^2 = .89$). The modulation index was larger in monophasic
485 than in biphasic coupling (13.98 vs. 9.96 ; $t(1,99) = 22.49$, $p < .01$, $\omega^2 = .83$; Figure 4A).
486 That is, multimodality influences the three methods very differently. Phase-locking value and mean
487 vector length cannot find biphasic coupling as it was implemented here (amplitude of the higher
488 frequency was increased at peak and trough of the lower frequency). Because of the mathematic
489 construct of the mean vector length (equation 2, Figure 1B) this is not surprising. Peak and trough appear
490 on opposite sides in the polar plane: their mean will cancel each other out. If other forms of biphasic
491 coupling would be present, the mean vector length could be able to find it, but would probably
492 underestimate its strength and would furthermore return distorted phase information. Therefore, it is
493 important to have a look at the polar plot before interpreting one's results. Similarly, the phase-locking
494 value cannot detect biphasic coupling, as it was implemented here. For biphasic coupling the amplitude
495 envelope oscillates twice as fast as the lower frequency band. Because of this, the phase lag between
496 lower and upper frequency band spans the entire polar plane. The modulation index is able to find
497 biphasic coupling, but biphasic coupling leads to a reduction in the phase-amplitude coupling value.
498 Literature indicates that biphasic coupling plays a minor role in empiric data. To our knowledge only a
499 very small fraction of studies report biphasic coupling (e. g. Lega et al., 2016, Leszczynski et al., 2015,
500 van der Meij et al., 2012). Most studies report monophasic coupling.

501

502 *3.2.5. Effect of data length on phase-amplitude coupling measures*

503 Coupling values of all methods increased with increasing data length (main effect data length: $F(2,198)$
504 $= 349.13$, $p < .01$, $\omega^2 = .70$). For the shortest epoch of 400 ms, none of the methods could detect
505 significant coupling, even though it was engineered into the data. The interaction method by data length
506 ($F(4,396) = 240.65$, $p < .01$, $\omega^2 = .52$; Figure 4B) became significant. Post hoc t-tests showed that all
507 factor levels within a method differed significantly from each other (all p 's $< .01$). The data length effect
508 was most pronounced for mean vector length ($.82 < \omega^2 < .94$), and phase-locking value ($.80 < \omega^2 < .94$).
509 The modulation index was least affected by data length ($.65 < \omega^2 < .75$).
510 Overall, the longer the data, the larger phase-locking value, mean vector length, and modulation index
511 are. This association was found in the data presented here, but must not generally apply. Here coupling

512 was simulated continuously into the data. If coupling is transient and does not proportionally vary with
513 data length, this relationship does not need to apply. Penny et al. (2008) showed, that coupling strength
514 decreases for phase-amplitude coupling, which was simulated transiently. Potentially, the general rule
515 is that the longer the data epochs where coupling occurs, the stronger the phase-amplitude coupling
516 values. This should be tested in a follow-up analysis. This analysis further showed that a minimal data
517 length is required for finding coupling, which should exceed at least 400 milliseconds per trial when
518 including 30 trials (also see Cheng et al., 2018). None of the methods were able to detect coupling in
519 the shortest simulated epoch of 400 milliseconds. It might be useful to develop a correction factor (e. g.
520 similar to the pairwise phase consistency that is insensitive to data length variation; Vinck et al., 2010)
521 for data length, to make phase-amplitude coupling values more comparable across studies. Of all three
522 methods, modulation index is least affected from the confounding factor data length.

523

524 3.2.6. Effect of sampling rate on phase-amplitude coupling measures

525 Overall coupling values slightly increased with increasing sampling rate ($F(1,99) = 23.65, p < .01, \omega^2 =$
526 $.10$). The sampling rate effect differed according to the method ($F(2,198) = 14.02, p < .01, \omega^2 = .04$;
527 Figure 4C). It was most pronounced in the mean vector length ($t(99) = -5.15, p < .01, \omega^2 = .20$), followed
528 by the phase-locking value ($t(99) = -4.86, p < .01, \omega^2 = .18$). The modulation index was least affected
529 by sampling rate ($t(99) = -4.23, p < .01, \omega^2 = .14$).

530 The factor sampling rate stands out because of its comparatively small effect size. A second set of data
531 was simulated testing phase-locking value, mean vector length, and modulation index at 16 – 18 Hz for
532 the modulating frequency and 202 – 238 Hz for the modulated frequency (for detailed results see
533 Hülsemann, 2016). This analysis showed that sampling rate is indeed important, but only if the
534 investigated upper frequency band approaches the Nyquist frequency. Of all three methods, modulation
535 index is least affected from the confounding factor sampling rate.

536

537 3.2.7. Effect of noise on phase-amplitude coupling measures

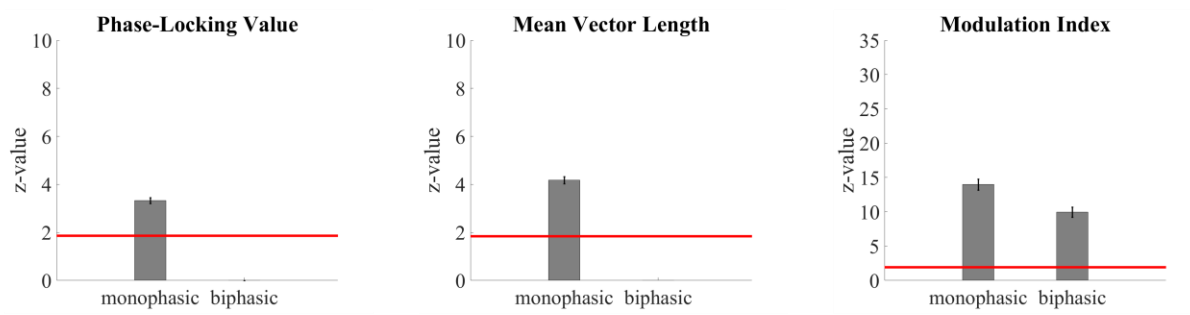
538 Coupling values of all methods decreased with increasing noise ($F(2,198) = 325.22, p < .01, \omega^2 = .68$).
539 The interaction method by noise became significant ($F(4,396) = 251.00, p < .01, \omega^2 = .53$; Figure 4D).

540 Post hoc t-tests showed that all factor levels within a method differed significantly from each other (all
541 p 's < .01). The effect of noise was most pronounced for the modulation index ($.65 < \omega^2 < .76$) and the
542 mean vector length ($.55 < \omega^2 < .84$). The phase-locking value was least affected by noise ($.51 < \omega^2 <$
543 $.81$).

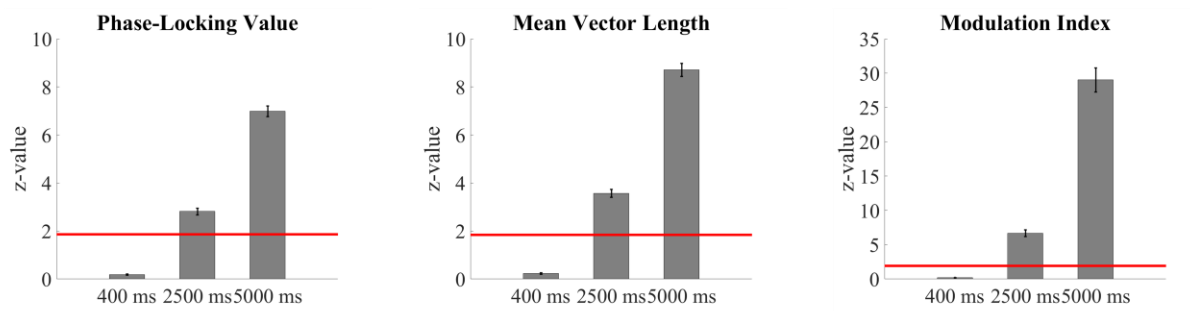
544 Overall, the noisier the data, the lower phase-locking value, mean vector length, and modulation index
545 are. This aspect is not desired but plausible. Noise obscures the relation between the phase of the lower
546 frequency and amplitude of the higher frequency. The data as a whole contains phase-amplitude
547 coupling to a lesser extent, as the relative amount of noise compared to the relative amount of signal
548 increases. Of all three methods, phase-locking value is least affected from the confounding factor noise.

549

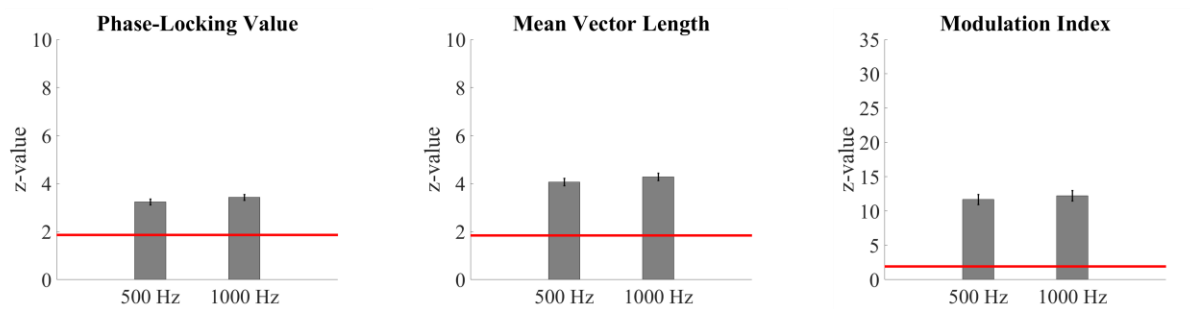
A) Multimodality Effect



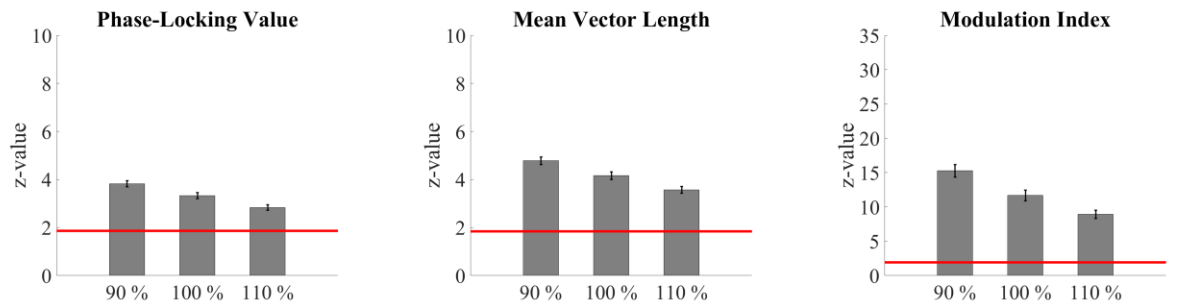
B) Data Length Effect



C) Sampling Rate Effect



D) Noise Effect



550

551

[2 column fitting image]

552

553

554

555

Figure 4 – Moderators of the phase-amplitude coupling measures: Mean (\pm SEM) phase-amplitude coupling values for each method for the A) multimodality effect, B) data length effect, C) sampling rate effect, and D) noise effect. In contrast to the modulation index, biphasic coupling could not be detected by the phase-locking value and mean vector length. This factor might turn out to be not as important, as

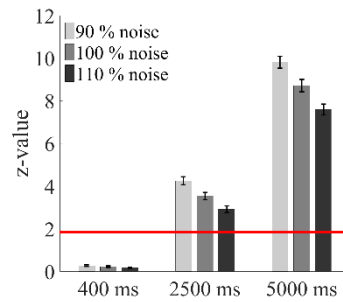
556 most studies report monophasic coupling. Coupling values of all methods increased with increasing data
557 length and slightly increase with sampling rate. Sampling rate only becomes relevant when analysing
558 frequencies close to the Nyquist frequency. Of all three methods, modulation index is least affected from
559 the confounding factor data length and sampling rate. Coupling values of all methods decreased with
560 increasing noise, while the phase-locking value is least affected from this confounding factor. The red
561 line marks the significance level. All values above this line represent significant phase-amplitude
562 coupling. For each effect, all factor levels within a method are significantly different from each other
563 according to post-hoc t-tests. For B), C), and D) only monophasic coupling values are depicted for the
564 phase-locking value and the mean vector length.

565

566 3.2.8. Interaction Effects

567 Conducting a 6-way ANOVAs for each method separately (see Hülsemann, 2016 for detailed results),
568 revealed ordinal interaction for all factors (multimodality, data length, sampling rate, noise, modulation
569 strength, and modulation width). Especially multimodality and data length interacted with the remaining
570 factors, as well as interacted with each other and the remaining factors. Sampling rate only showed
571 significant interactions, when analysing frequencies close to the Nyquist frequency. All interactions had
572 a monotone pattern, following the pattern of each main effect. For example, mean vector length
573 increased the longer the data, but it increased less when also noise increases (Figure 5). This pattern was
574 true for each added factor. Phase-locking value and mean vector length did not find biphasic coupling
575 at all. Because of this, for these two methods, the described main effect and interaction patterns are only
576 valid for monophasic, but not for biphasic coupling. For the modulation index the pattern was true for
577 mono- and for biphasic coupling.

578



579

580

[1 column fitting image]

581 **Figure 5 – Interaction effects between the moderators of the phase-amplitude coupling measures:**

582 Mean (\pm SEM) phase-amplitude coupling values for the mean vector length for the data length by noise

583 interaction (only monophasic coupling values). Interactions had a monotone pattern, following the

584 pattern of each main effect. Depicted here, mean vector length increased the longer the data, but it

585 increased less when also noise increased. This pattern was true for each added factor. The red line marks

586 the significance level. All values above this line represent significant phase-amplitude coupling. For

587 each method, all factor levels are significantly different from each other according to Dunn's post hoc

588 test. Only values within the 400 ms condition do not differ between the noise levels.

589

590 Comparing all three methods it becomes evident that the modulation index is least affected by the

591 confounding factors multimodality, data length and sampling rate. However, it is also – like the phase-

592 locking value – less sensitive to variation in modulations strength compared with the mean vector length.

593 The modulation index is especially less sensitive to modulation width compared to the mean vector

594 length and phase-locking value. Mean vector length and modulation index are similarly – and stronger

595 than the phase-locking value – affected by the confounding factor noise.

596

597

598 **4. Conclusion**

599 In conclusion, for long data epochs, recorded at high sampling rates, with a high signal-to-noise ratio,

600 the use of the mean vector length is recommended, because it is more sensitive to modulation strength

601 and width than both other methods. For noisier data, shorter data epochs, recorded at a lower sampling

602 rate, the use of the modulation index is recommended, as it is least influenced by the confounding factors

603 compared with both other methods. If it is not clear whether cross-frequency coupling will be mono- or
604 biphasic, the modulation index should be used, even though literature suggests that biphasic coupling
605 can be neglected.

606 The phase-locking value does not stand out in comparison to the two other measures. Its usage is
607 potentially problematic because phase information is extracted from the amplitude envelope of a signal.
608 Phase information can only be correctly extracted from truly oscillating signals; this must not be
609 necessarily the case for an amplitude envelope. So far, no review evaluated this measure explicitly as
610 positive.

611 Because mean vector length and modulation index have complementing strengths and weaknesses, it
612 would be advisably to calculate both. The time-consuming aspect of measuring phase-amplitude
613 coupling is permutation testing. Calculation of both measures on the other hand will not substantially
614 increase the analysis time.

615 The modulation index is quantitatively larger than the phase-locking value and mean vector length.
616 However, even despite substantial quantitative differences in values, the qualitative decision for
617 significance of phase-amplitude coupling is the same for all three methods in our simulation.
618 Nevertheless, comparison of coupling strengths between the methods is problematic and this lack of
619 comparability provides another reason for reporting both, mean vector length and modulation index.

620 In contrast to mean vector length, the false positive rate of the modulation index is not affected by any
621 confounding factor. However, this advantage against mean vector length is counteracted by one
622 disadvantage against the mean vector length: calculation of the modulation index includes Shannon's
623 Entropy. The entropy value depends on the amount of bins as well as amount of data squeezed into the
624 same amount of bins. This is an undesirable degree of freedom, which is not present when calculating
625 the mean vector length.

626 Due to the dependency on confounding variables (e. g. data length), comparing absolute coupling
627 strengths across studies might be difficult even if using the same method. Comparisons within one study,
628 on the other hand, can be done with confidence. Nevertheless, one should make sure that signal-to-noise
629 ratio is comparable within all experimental conditions and over the course of the experiment.

630 Generally, it is advisable to work with standardized phase-amplitude coupling measures via permutation
631 testing. It facilitates the interpretation of the measures, first and foremost, by giving the researcher
632 knowledge about the probability that the observed modulation index would have been also found under
633 the assumption of the null-hypothesis. This aspect is often ignored in the literature.
634 Kramer and Eden (2013) stated that “an optimal analysis method to assess this cross-frequency coupling
635 (CFC) does not yet exist” (p.64). Even if it would be ideal, to have a measure that is less susceptible to
636 confounding variables summarizing this analysis, it should be rather concluded that at least two
637 reasonable analysis methods exist.

638

639 **Acknowledgements: -**

640 References

- 641 Berman, J.I., McDaniel, J., Liu, S., Cornew, L., Gaetz, W., Roberts, T.P., Edgar, J.C., 2012. Variable
642 bandwidth filtering for improved sensitivity of cross-frequency coupling metrics. *Brain*
643 *connectivity* 2 (3), 155–163. 10.1089/brain.2012.0085.
- 644 Bruns, A., Eckhorn, R., 2004. Task-related coupling from high- to low-frequency signals among visual
645 cortical areas in human subdural recordings. *International Journal of Psychophysiology* 51 (2),
646 97–116. 10.1016/j.ijpsycho.2003.07.001.
- 647 Burgess, A.P., Ali, L., 2002. Functional connectivity of gamma EEG activity is modulated at low
648 frequency during conscious recollection. *International journal of psychophysiology : official*
649 *journal of the International Organization of Psychophysiology* 46 (2), 91–100. 10.1016/S0167-
650 8760(02)00108-3.
- 651 Canolty, R.T., Edwards, E., Dalal, S.S., Soltani, M., Nagarajan, S.S., Kirsch, H.E., Berger, M.S., Barbaro,
652 N.M., Knight, R.T., 2006. High gamma power is phase-locked to theta oscillations in human
653 neocortex. *Science (New York, N.Y.)* 313 (5793), 1626–1628. 10.1126/science.1128115.
- 654 Cheng, N., Li, Q., Wang, S., Wang, R., Zhang, T., 2018. Permutation Mutual Information: A Novel
655 Approach for Measuring Neuronal Phase-Amplitude Coupling. *Brain topography* 31 (2), 186–201.
656 10.1007/s10548-017-0599-2.
- 657 Cohen, M.X., 2008. Assessing transient cross-frequency coupling in EEG data. *Journal of Neuroscience*
658 *Methods* 168 (2), 494–499. 10.1016/j.jneumeth.2007.10.012.
- 659 Cohen, M.X., 2014. *Analyzing neural time series data: Theory and practice*. The MIT Press, Cambridge,
660 Massachusetts, 1 online resource (xviii, 578).
- 661 Dunn, O.J., 1961. Multiple Comparisons among Means. *Journal of the American Statistical*
662 *Association* 56 (293), 52–64. 10.1080/01621459.1961.10482090.
- 663 Dvorak, D., Fenton, A.A., 2014. Toward a proper estimation of phase-amplitude coupling in neural
664 oscillations. *Journal of Neuroscience Methods* 225, 42–56. 10.1016/j.jneumeth.2014.01.002.
- 665 Geisser, S., Greenhouse, S.W., 1958. An Extension of Box's Results on the Use of the F Distribution in
666 Multivariate Analysis. *The Annals of Mathematical Statistics* 29 (3), 885–891.
667 10.1214/aoms/1177706545.
- 668 Hays, W.L., 1973. *Statistics for the social sciences*, 2nd ed. Holt, Rinehart and Winston, New York, xxi,
669 954.
- 670 He, B.J., Zempel, J.M., Snyder, A.Z., Raichle, M.E., 2010. The temporal structures and functional
671 significance of scale-free brain activity. *Neuron* 66 (3), 353–369. 10.1016/j.neuron.2010.04.020.
- 672 Hülsemann, M.J., 2016. *The Role of Phase-Amplitude Coupling in the Relationship between Acute*
673 *Stress and Executive Functions*, Universitätsring 15, 54296 Trier.
- 674 Jensen, O., 2006. Maintenance of multiple working memory items by temporal segmentation.
675 *Neuroscience* 139 (1), 237–249. 10.1016/j.neuroscience.2005.06.004.
- 676 Jensen, O., Lisman, J.E., 1998. An oscillatory short-term memory buffer model can account for data
677 on the Sternberg task. *The Journal of neuroscience : the official journal of the Society for*
678 *Neuroscience* 18 (24), 10688–10699.
- 679 Kramer, M.A., Eden, U.T., 2013. Assessment of cross-frequency coupling with confidence using
680 generalized linear models. *Journal of Neuroscience Methods* 220 (1), 64–74.
681 10.1016/j.jneumeth.2013.08.006.
- 682 Lakatos, P., Shah, A.S., Knuth, K.H., Ulbert, I., Karmos, G., Schroeder, C.E., 2005. An oscillatory
683 hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal*
684 *of neurophysiology* 94 (3), 1904–1911. 10.1152/jn.00263.2005.
- 685 Lega, B., Burke, J., Jacobs, J., Kahana, M.J., 2016. Slow-Theta-to-Gamma Phase-Amplitude Coupling in
686 Human Hippocampus Supports the Formation of New Episodic Memories. *Cerebral cortex (New*
687 *York, N.Y. : 1991)* 26 (1), 268–278. 10.1093/cercor/bhu232.
- 688 Leszczynski, M., Fell, J., Axmacher, N., 2015. Rhythmic Working Memory Activation in the Human
689 Hippocampus. *Cell reports* 13 (6), 1272–1282. 10.1016/j.celrep.2015.09.081.

- 690 Lisman, J.E., Jensen, O., 2013. The theta-gamma neural code. *Neuron* 77 (6), 1002–1016.
691 10.1016/j.neuron.2013.03.007.
- 692 Luck, S.J., 2014. *An Introduction to the Event-Related Potential Technique*, 2nd ed. The MIT Press,
693 Cambridge, 1417 pp.
- 694 Maris, E., van Vugt, M., Kahana, M., 2011. Spatially distributed patterns of oscillatory coupling
695 between high-frequency amplitudes and low-frequency phases in human iEEG. *NeuroImage* 54
696 (2), 836–850. 10.1016/j.neuroimage.2010.09.029.
- 697 Miller, K.J., Sorensen, L.B., Ojemann, J.G., den Nijs, M., 2009. Power-law scaling in the brain surface
698 electric potential. *PLoS computational biology* 5 (12), e1000609. 10.1371/journal.pcbi.1000609.
- 699 Mormann, F., Fell, J., Axmacher, N., Weber, B., Lehnertz, K., Elger, C.E., Fernandez, G., 2005.
700 Phase/amplitude reset and theta-gamma interaction in the human medial temporal lobe during a
701 continuous word recognition memory task. *Hippocampus* 15 (7), 890–900. 10.1002/hipo.20117.
- 702 Novak, P., Lepicovska, V., Dostalek, C., 1992. Periodic amplitude modulation of EEG. *Neuroscience*
703 *letters* 136 (2), 213–215. 10.1016/0304-3940(92)90051-8.
- 704 Onslow, A.C.E., Bogacz, R., Jones, M.W., 2011. Quantifying phase-amplitude coupling in neuronal
705 network oscillations. *Progress in biophysics and molecular biology* 105 (1-2), 49–57.
706 10.1016/j.pbiomolbio.2010.09.007.
- 707 Özkurt, T.E., Schnitzler, A., 2011. A critical note on the definition of phase-amplitude cross-frequency
708 coupling. *Journal of Neuroscience Methods* 201 (2), 438–443. 10.1016/j.jneumeth.2011.08.014.
- 709 Penny, W.D., Duzel, E., Miller, K.J., Ojemann, J.G., 2008. Testing for nested oscillation. *Journal of*
710 *Neuroscience Methods* 174 (1), 50–61. 10.1016/j.jneumeth.2008.06.035.
- 711 Pfurtscheller, G., 1976. Ultralangsame Schwankungen innerhalb der rhythmischen Aktivität im Alpha-
712 Band und deren mögliche Ursachen. *Pflügers Arch. (Pflügers Archiv European Journal of*
713 *Physiology)* 367 (1), 55–66. 10.1007/BF00583657.
- 714 Rasch, B., Frieze, M., Hofmann, W., Naumann, E., 2006. *Quantitative Methoden 2: Einführung in die*
715 *Statistik*, 2nd ed. Springer Medizin Verlag Heidelberg, Berlin, Heidelberg, 1 online resource
716 (Online-Ressource.).
- 717 Tort, A.B.L., Komorowski, R., Eichenbaum, H., Kopell, N., 2010. Measuring phase-amplitude coupling
718 between neuronal oscillations of different frequencies. *Journal of neurophysiology* 104 (2), 1195–
719 1210. 10.1152/jn.00106.2010.
- 720 Tort, A.B.L., Kramer, M.A., Thorn, C., Gibson, D.J., Kubota, Y., Graybiel, A.M., Kopell, N.J., 2008.
721 Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and
722 hippocampus during performance of a T-maze task. *Proceedings of the National Academy of*
723 *Sciences of the United States of America* 105 (51), 20517–20522. 10.1073/pnas.0810524105.
- 724 van der Meij, R., Kahana, M., Maris, E., 2012. Phase-amplitude coupling in human
725 electrocorticography is spatially distributed and phase diverse. *The Journal of neuroscience : the*
726 *official journal of the Society for Neuroscience* 32 (1), 111–123. 10.1523/JNEUROSCI.4816-
727 11.2012.
- 728 Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., Pennartz, C.M.A., 2010. The pairwise phase
729 consistency: A bias-free measure of rhythmic neuronal synchronization. *NeuroImage* 51 (1), 112–
730 122. 10.1016/j.neuroimage.2010.01.073.
- 731 Vosskuhl, J., Huster, R.J., Herrmann, C.S., 2015. Increase in short-term memory capacity induced by
732 down-regulating individual theta frequency via transcranial alternating current stimulation.
733 *Frontiers in human neuroscience* 9, 257. 10.3389/fnhum.2015.00257.
- 734 Widmann, A., Schröger, E., Maess, B., 2015. Digital filter design for electrophysiological data – a
735 practical approach. *Journal of Neuroscience Methods* 250, 34–46.
736 10.1016/j.jneumeth.2014.08.002.
- 737 Zhivomirov, H., 2013. Pink, Red, Blue and Violet Noise Generation with Matlab Implementation.
738 [http://www.mathworks.com/matlabcentral/fileexchange/42919-pink--red--blue-and-violet-](http://www.mathworks.com/matlabcentral/fileexchange/42919-pink--red--blue-and-violet-noise-generation-with-matlab-implementation/content/rednoise.m)
739 [noise-generation-with-matlab-implementation/content/rednoise.m](http://www.mathworks.com/matlabcentral/fileexchange/42919-pink--red--blue-and-violet-noise-generation-with-matlab-implementation/content/rednoise.m). Accessed 17 April 2016.
- 740