Electroencephalography and clinical Neurophysiology, 1983, 55: 468–484
Elsevier Scientific Publishers Ireland, Ltd.

Technical Section

A NEW METHOD FOR OFF-LINE REMOVAL OF OCULAR ARTIFACT [1]

GABRIELE GRATTON, MICHAEL G.H. COLES and EMANUEL DONCHIN

Cognitive Psychophysiology Laboratory, Department of Psychology, University of Illinois, Champaign, Ill. 61820 (U.S.A.)

(Accepted for publication: November 24, 1982)

Ocular potentials are clearly a nuisance in investigations of event-related brain potentials (ERPs). Eye movements, or blinks, that occur in synchrony with the events eliciting the ERP introduce unwelcome artifacts into the data. This is especially the case because the frequency spectra of endogenous ERP components (such as the CNV and P300) are similar to the spectra of the oculographic potentials (e.g., Hillyard and Galambos 1970; Girton and Kamiya 1973). Treating oculographic artifacts has therefore become a required procedure in ERP research (Donchin et al. 1977).

Movements of the eye and of the eyelid generate changes in electrical fields that interact with fields generated by intracranial generators and thus affect the electrical activity at the scalp. Two different mechanisms have been proposed to account for these ocular potentials. First, rotation of the eyeball modifies scalp recorded electrical activity because movement of the dipole between cornea and retina results in electrical field changes which are propagated through the skull (Mowrer et al. 1935; Overton and Shagass 1969). Second, the upper and lower eyelids act as 'sliding electrodes' as they move across the eyeball creating an electrical field which can also be propagated through the skull (Barry and Jones 1965).

The procedure commonly used to eliminate ocular artifacts requires the investigator to detect and discard data recorded on trials in which excessive eye movements, or blinks, occur. To this end, ocular activity is monitored using the electro-oculogram (EOG). All trials are rejected for which some attribute of the EOG trace (such as amplitude, variance, area, or slope) exceeds a criterion value.

Unfortunately, this simple procedure is somewhat problematic. First, eye movements and blinks may have important effects on those perceptual and cognitive processes that are under investigation (Carpenter 1948; Drew 1951). Thus, rejection of trials associated with ocular activity may lead to the use of unrepresentative samples of trials in the computation of the average ERP. Second, certain classes of subjects cannot, or will not, cooperate and avoid ocular movements during the trials. In these cases (e.g., children and psychiatric or neurological patients) it is difficult, if not impossible, to obtain a sufficient sample of trials free of ocular movement. Third, some tasks of interest to ERP researchers involve complex tracking or scanning of the visual field. In these tasks, ocular activity is inherent and therefore the proportion of 'clean' trials is likely to be small indeed. Further, these 'clean' trials are, by definition, unrepresentative.

The trial rejection procedure is often accompanied by instructions to the subjects to 'keep their eyes still' during the experiment, or to confine blinks and movements to designated times when experimental stimuli are not being presented. These instructions effectively assign a 'secondary task' to the subject, requiring the division of resources

[1] This research reported here has been supported under the Office of Naval Research (Contract No. N00014-76-C-0002) with funds provided by the Defense Advanced Research Projects Agency; AFOSR under Contract No. F49620-79C-0233; Wright Patterson AFB under Contract No. F33615-79C-0512; and the Environmental Protection Agency under Contract No. R805628010.

0013-4649/83/0000-0000/$03.00

between the experimental task and self-monitoring of ocular activity. Note that, if there are individual differences in ocular activity, the secondary task will vary in difficulty from subject to subject.

Alternative procedures seek to correct the ERP data for the effect of ocular activity. These procedures assume that the scalp potential is given by the linear summation of brain and ocular potentials. Therefore, if we subtract the ocular potentials from the EEG record, we should obtain a 'clean' brain potential. According to the simplest version of this procedure, the EOG is subtracted from each trial or from the average (either on-line or off-line) to remove that portion of the raw EEG that can be attributed to ocular activity. As the EOG signal is presumed to propagate by volume conduction across the skull, it is reasonable to expect that the contribution of EOG varies across the scalp. Thus, subtraction of the *same* EOG signal at each electrode site is likely to yield erroneous estimates of the residual ERP. A more reasonable approach would attempt to estimate a propagation factor for each electrode site and scale the EOG signal by that factor prior to subtraction.

There have been several attempts to estimate these propagation factors. Hillyard and Galambos (1970) and Girton and Kamiya (1973) estimated the propagation factor by recording the EOG elicited when subjects moved their eyes by known amounts during a pre-experimental calibration session. The effects of these movements at different electrode locations were used to estimate specific correction factors, and regression techniques (Hillyard and Galambos 1970) or analog procedures (Girton and Kamiya 1973) were used to derive corrected values of EEG. This technique, then, depends for its validity both on the linearity assumptions and on the validity of the estimation procedure for the propagation factor.

The assumption of linearity is supported by data obtained by Hillyard and Galambos (1970) and by Overton and Shagass (1969). Slight deviation from linearity is found for vertical movements, due probably to differential involvement of the eyelids in upward versus downward movements (Overton and Shagass 1969). But these deviations are not sufficient to violate the assumption of linearity.

However, the assumptions underlying the estimation of the correction factor may be untenable. Thus, it was assumed that the fields generated by voluntary movements made during a pre-experimental calibration session are identical to the fields generated by task contingent eye movements. There is evidence that voluntary and involuntary eye movements are different (Records 1979). These differences may lead to differences in the nature of the ocular artifact. It would seem prudent to estimate the propagation factor under the *same* conditions that obtain when it is applied. A second assumption is that fields generated by blinks and eye movements are the same. Overton and Shagass (1969) have demonstrated convincingly that this is not the case. While both blinks and eye movements affect the ocular dipole, there is considerable eyelid activity during blinks which creates characteristic fields. Thus, a correction factor for blinks should be estimated separately from a correction factor due to eye movements. The procedure recently proposed by Verleger et al. (1982) involves the computation of propagation factors from a subset of the data which needs correction. In this sense, it avoids the first problem mentioned above. However, the procedure fails to differentiate between blinks and saccades and, thus, may yield invalid correction factors.

We describe here an eye movement correction procedure (EMCP) for computing the correction factors that avoids both problems. First, estimates of correction factors are derived from EOG and EEG records obtained during the experiment rather than during a pre-experimental calibration period. Second, correction factors are computed separately for blinks and eye movements. A third and novel feature of the current procedure is that the degree to which signals at the EOG electrodes are propagated to any given EEG electrode pair is estimated after removal of *event-related* EOG and EEG activity from the data. If there are event-related effects in both EEG and EOG records, the propagation of EOG signal to EEG electrode would be overestimated. Our estimates of correction factors are computed on data at each time point on every trial *after* event-related activity has been subtracted.

After a description of the algorithm we present

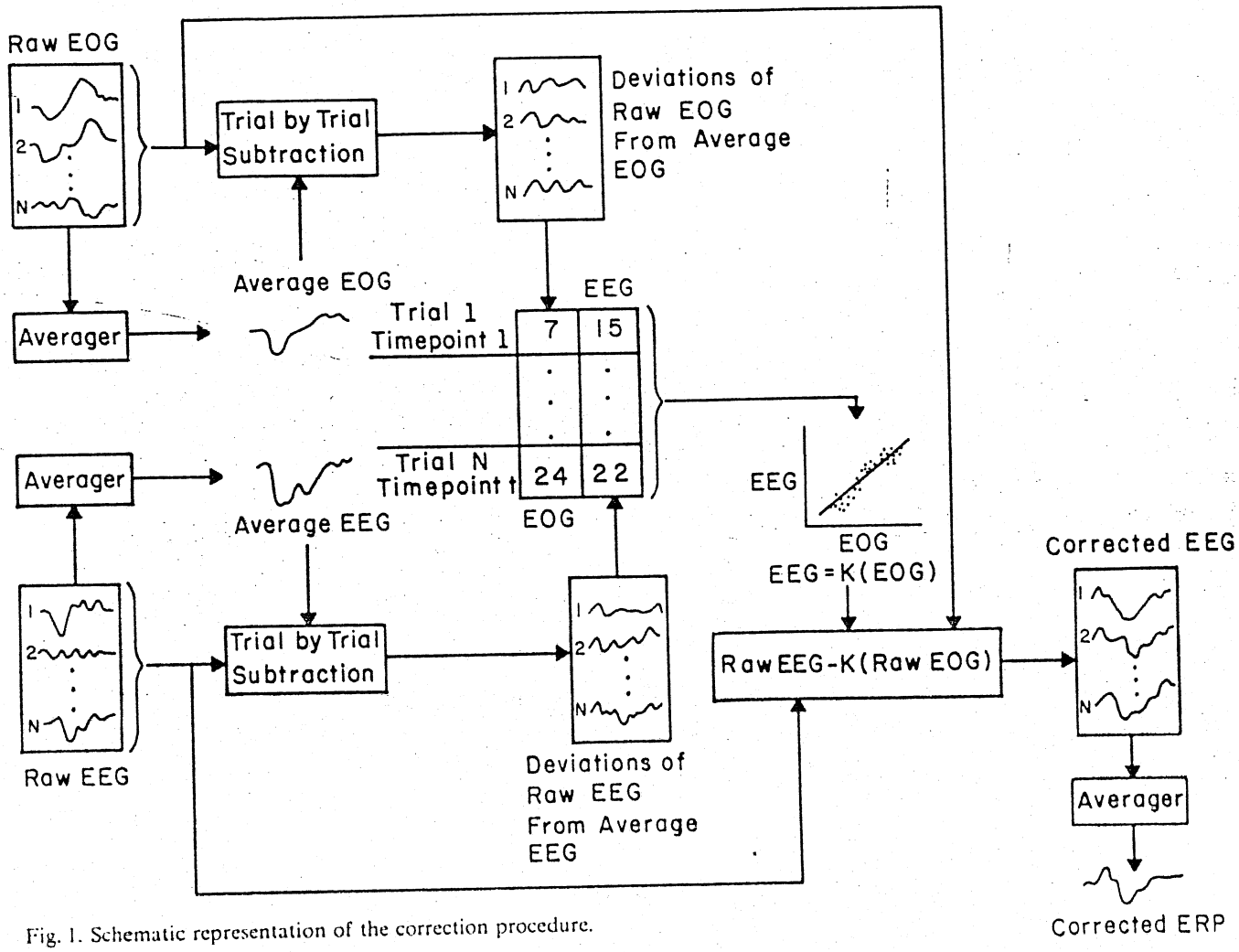

Fig. 1. Schematic representation of the correction procedure.

several tests that confirm the validity and reliability of this off-line procedure for removing ocular artifacts from the EEG.

Procedure

We begin by describing the computational procedures that are used to derive the correction factors. A formal derivation is presented in Appendix A. Fig. 1 gives a graphic representation of these procedures. Assume, for the sake of this description, that a subject is run in a study with P conditions and that N trials, and thus N records, are available for each condition. The record for each trial consists of data, obtained at K electrode pairs, one of which abuts the eyeball yielding an EOG trace for each of the trials [2].

[2] If primary interest is in ERPs from mid-line derivations only vertical EOG need be obtained. If interest is also in ERPs from lateral electrode sites, then both horizontal and vertical EOG should be obtained and the correction procedure applied separately for each EOG channel. Picton (personal communication) indicated that spuriously high correction factors, for either horizontal or vertical movements, could be obtained if there are consistent oblique movements. A solution to this problem is to detect time points in which such movements occur using vertical and horizontal EOG, and then either compute a separate correction factor for these movements, or discard these time points from the data bases used to compute vertical or horizontal correction factors. It should be noted that correction for blinks is accomplished using vertical EOG only. EEG and EOG would be recorded using the same filtering parameters.

(a) Raw averaging

For each of the P conditions all the N trials are averaged, for each of the K electrode pairs, separately. These ensemble averages estimate the stimulus related variation for the EEG and the EOG channels.

(b) Subtraction of the raw average from the single trial data

The raw averages are subtracted from the corresponding single trial records. That is, the average is subtracted from each of the N trials used in its own computation. After subtraction of a 'baseline' computed as the mean of all points in an epoch, we have a set of wave forms that are considered an estimate of the activity at an electrode site, on each trial, that is *not* event-related.

Correction factor

Step (b) yields estimates of the electrical activity for each trial at the EOG and the EEG electrodes that is not stimulus related. The EEG data are then considered as a dependent variable in a regression computation in which the EOG data serve as independent variable. The regression equations computed in this fashion are used to derive correction factors.

Separation of blinks and eye movements

Since correction factors are computed separately for blinks and eye movement data. Blinks are selected first by locating all time points in the EOG record where the rate of change in electrical activity exceeds a criterion. These points are used to form a correction factor as described above. The (pre-subtraction) EEG data are then 'corrected' by subtracting from the EEG, the EOG multiplied by the correction factor. A new correction factor for eye movements is then computed from the data remaining after the blinks have been removed. Note that although the correction factor is computed using only activity that is *not* blinks, it is applied to *all* activity in the

Averaging

After subtracting the scaled EOG from all single trials, the data can be averaged as in (a) yielding corrected ERPs.

Tests of the procedure

Three series of tests have been conducted to evaluate the validity and the reliability of EMCP.

Series 1

The first series of tests involved data obtained from a study on the relationship between P300, slow wave and CNV (Donchin et al. 1983). Subjects were required to perform a simulated driving task in which they controlled a 'car' with a joystick as it appeared to move along a road depicted in a CRT display. From time to time, numbers appeared within the subject's car which warned the subject of the upcoming appearance of obstacles which had to be avoided. EEG measures, obtained from Fz, Cz and Pz, were used to derive ERP responses to the warning numbers. EOG measures [3] were also obtained from the same epoch (1400 msec). For each of the subjects, data obtained on about 100 trials were used. The nature of the task and the display we used led to a sufficient number of trials being associated with eye movements, providing a reasonable data base for these validity tests. Two types of tests are presented: (A) deviation from 'true' ERP, and (B) reduction in variance.

For the purpose of these tests we note the following terminology:

(a) True ERP

For a given subject the average of the records obtained for trials during which EOG activity did not exceed a strict variance criterion.

(b) Dependent and independent samples

Two samples of trials were used to provide the ERPs to be compared with the 'true' ERP.

(i) A sample consisting of the trials excluded from the sample used to compute the 'true' ERP.

[3] An oblique electrode derivation was used to derive the EOG data in this experiment. Thus, variations in the EOG trace could be due to either vertical or horizontal movements (or both). In fact, task requirements were such that vertical (and not horizontal) movements occurred. Therefore, the EOG trace reflected predominantly vertical eye movement activity.

Thus, this sample contained all trials for which EOG activity exceeded our strict criterion. Because the sample contains a *different* set of trials than that used to compute the 'true' ERP, we use the term '*independent*' sample.

(ii) A sample consisting of *all* the trials. Because this sample included those trials used to compute the 'true' ERP, we refer to it as the '*dependent*' sample.

(c) Raw ERP

For a given subject, the raw ERP is defined as the average of *all* trials in either the independent or dependent samples. The records are not corrected and are used in averaging regardless of the occurrence of EOG artifacts.

(d) Corrected ERP

For a given subject, the corrected ERP is defined as the average of all trials in either the independent or dependent samples after the trials have been corrected by EMCP.

(e) Randomly corrected ERP

This ERP is computed in the same manner as the corrected ERP except that the correction factor is obtained by substituting a random number between −1 and +1 (from a rectangular distribution) for the correlation coefficient used in the correction procedure. In this way, 10 'random regression coefficients' were generated by multiplying these randomly selected values by the ratio of the standard deviations of EEG and EOG. This random-correction procedure was used to estimate a sampling distribution of deviations within which we could evaluate deviations yielded by the EMCP.

(A) Deviation from 'true' ERP. The first tests of validity of EMCP involve comparisons of corrected ERPs with an estimate of the 'true' ERP. While we had no direct measure of the 'true' ERP, we assumed that the ERP obtained from trials selected by the traditional rejection procedure can serve as a reasonable substitute. This appears plausible because this type of ERP is accepted as a 'clean' ERP by most investigators. If we can show that the ERP corrected by EMCP is more similar to this 'true' ERP than uncorrected or randomly corrected ERPs, then we can consider our procedure to be valid.

The degree to which an ERP corresponds to (or differs from) the 'true' ERP is expressed by means of a deviation index, computed as follows:

$$\text{Deviation} = \sqrt{\frac{\sum_{t=1}^{n} (X_t - ERP'_t)^2}{n}}$$

where X_t = value of the evaluated ERP at time t; ERP'_t = value of 'true' ERP at point t; n = number of points in an epoch.

In the first test of this series, the deviation index from the true ERP was computed for the corrected ERP, the raw ERP, and each of 10 randomly corrected ERPs for each of 5 subjects, for each electrode.

Figs. 2 and 3 show the values of these deviation indices. Fig. 2 shows results for the *independent* trial sample; Fig. 3 gives data for the *dependent* trial sample. The deviation for the raw ERP is marked by an asterisk, the deviation for the corrected ERP with an arrow and that for the 10 randomly corrected ERPs with bars. As a combined measure of total deviation across the 3 electrodes, we computed the square root of the mean of the squared individual deviations. Values for total deviation are also shown in Figs. 2 and 3.

It can be seen that deviations for randomly corrected ERPs are generally larger than those for ERPs obtained using the derived correction factor. For both independent and dependent trial samples, the deviation for the corrected ERPs is always smaller than the deviation for the raw ERP. In a few cases randomly corrected ERPs yield smaller deviations, but these can be explained as follows. First, for subject 4, the EOG trace associated with the 'true' ERP showed a small but consistent EOG response. Thus, the estimate of the 'true' ERP for this subject may not be a good estimate. Secondly, subject 2 showed little evidence of eye movements. Third, subject 1 had very few trials for which EOG variance was less than the strict criterion. Thus, the estimate of the 'true' ERP for this subject may have been contaminated with noise.

In Fig. 4 we display ERP wave forms for two subjects based on independent trial samples. Of particular interest are those for the 'true' ERP and those for corrected and uncorrected ERPs based

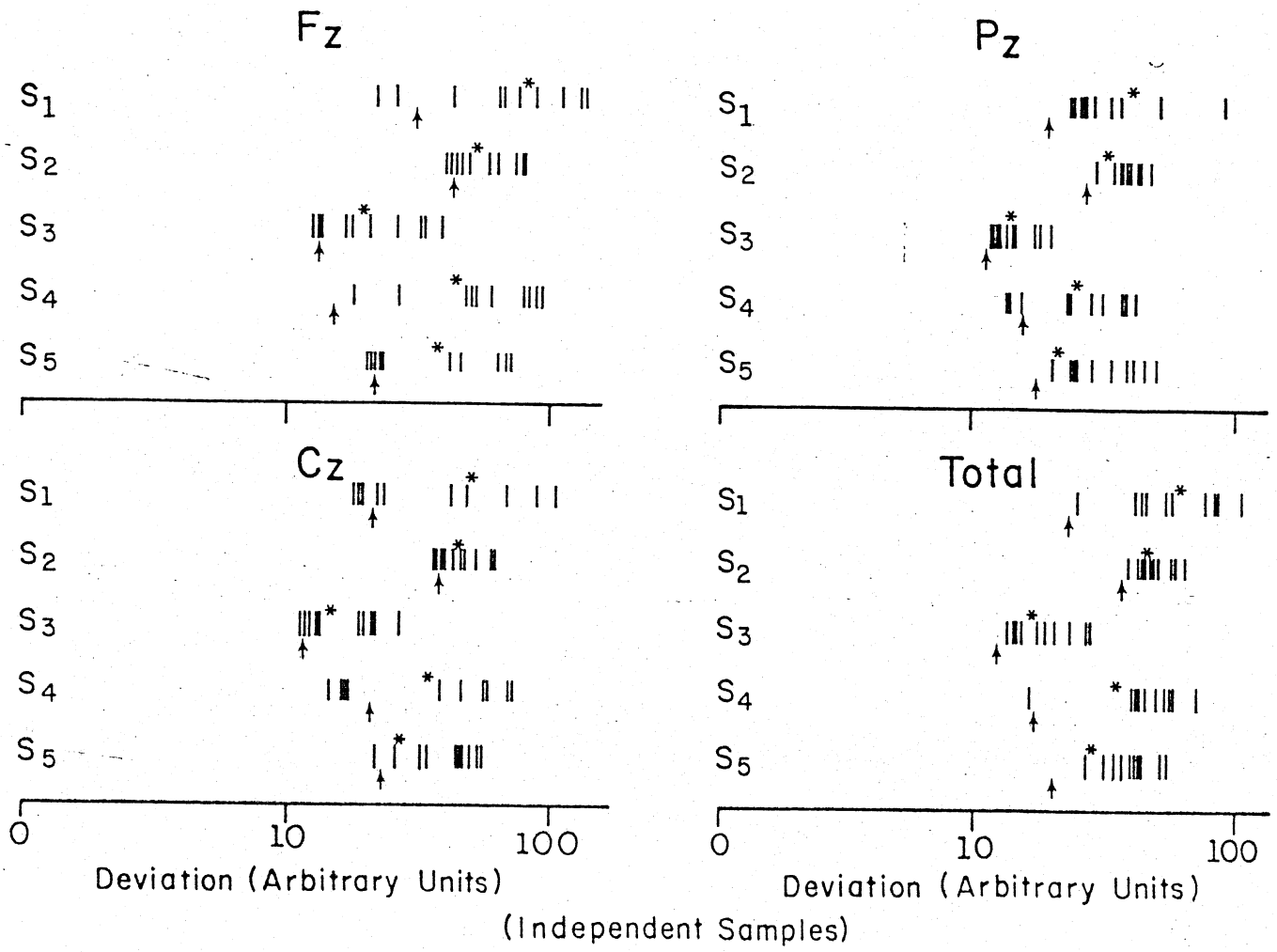

Fig. 2. Deviations (in log arbitrary units) from 'true' ERP of ERPs derived from trials with EOG variance larger than a criterion (independent sample). Deviations relative to wave forms corrected by EMCP (arrows), uncorrected (stars), and corrected according to a random procedure (bars) are shown separately for each subject and electrode. The total deviation computed over all electrodes is also shown.

on trials with EOG variance greater than the criterion. Note that, for both subjects, the similarity between corrected and 'true' ERPs is greater than that for raw and 'true' ERPs. Fig. 5 shows similar wave forms for subject 1 for the total trial sample. Note that, again, the effect of correction is to make the corrected ERP more similar to the 'true' ERP.

(B) Reduction in variance. We assumed above that the true ERP may be estimated by computing an average ERP based on trials selected for low EOG variance. This assumption, although made by almost every ERP researcher, may not be valid. Either because of residual consistent EOG activity, or because of inadequate sample size, the estimate of the true ERP may be inaccurate. For this reason, we conducted two additional validity tests that were based on the prediction that our correction procedure should reduce the variance between ERPs, or among trials at each time point. This is because some of the variance between trials is due to eye movements. The first of the additional tests examines the prediction that if EMCP increases the similarity between the corrected and the 'true' ERP, then corrected ERPs, derived from the same subject, condition and electrode, should be more alike than corresponding uncorrected ERPs.

The samples used in this test were the same as

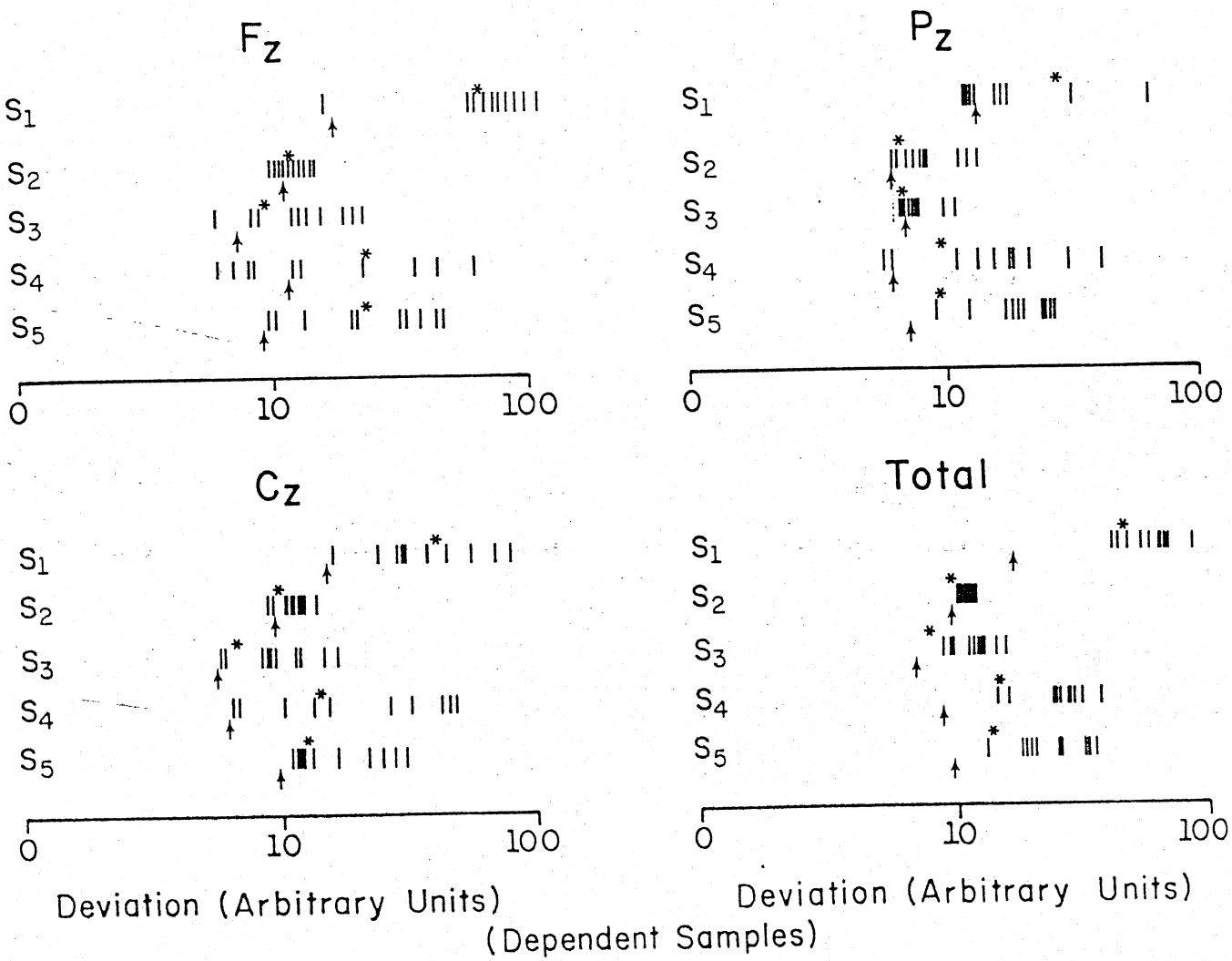

Fig. 3. Deviations (in log arbitrary units) from 'true' ERP of ERPs derived from all trials (dependent sample). Deviations relative to wave forms corrected by EMCP (arrows), uncorrected (stars) and corrected according to a random procedure (bars) are shown separately for each subject and electrode. The total deviation computed over all electrodes is also shown.

the *independent* samples used above. That is, for each subject, two sets of trials were derived, those for which EOG variance was greater, or less, than the strict criterion. Comparison between average ERPs based on these two samples, both before and after EMCP, was accomplished using the deviation index. Table I gives the relevant deviation indices. Note that, in every case, corrected ERPs are more similar to each other than are the raw ERPs. This is also illustrated in Fig. 4.

The second additional test examines the prediction that, if the values at each time point in the trial correspond more to those associated with the 'true' ERP as a result of correction, then the variance across trials for each time point should be reduced by correction. To test this prediction we computed, for each subject, electrode and time point, the variance across trials before and after correction. Then, differences in variance between uncorrected and corrected single trials were derived for each time point [4]. The frequency distributions of these differences are shown in Fig. 6.

For 2 subjects (out of 5) all the time points for all the electrodes show a reduction in the variance

[4] Note that, for each trial, the baseline, defined as the mean value of the time points across the epoch, is subtracted from each time point *before* the variance across trials is computed.

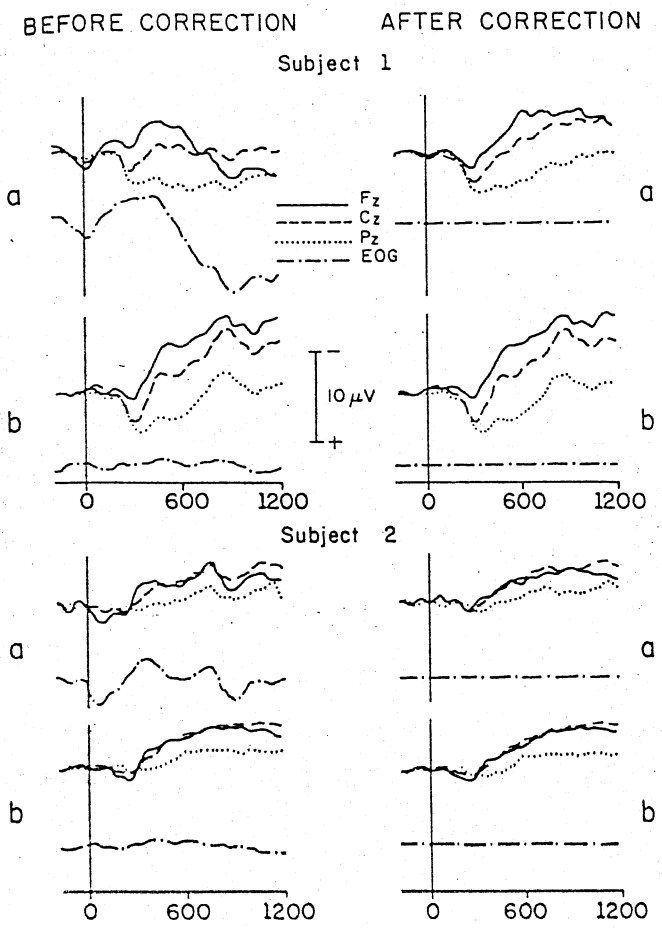

Fig. 4. ERPs and associated EOG for subjects 1 and 2 for trials with EOG variance larger (a) and smaller (b) than the criterion. Note that (a) corresponds to the 'independent' sample and (b) to the 'true' ERP sample. The figure shows the wave forms obtained from trials before correction (left) and after correction by EMCP (right). Note the ERP derived from uncorrected trials with EOG variance smaller than the criterion corresponds to the 'true' ERP.

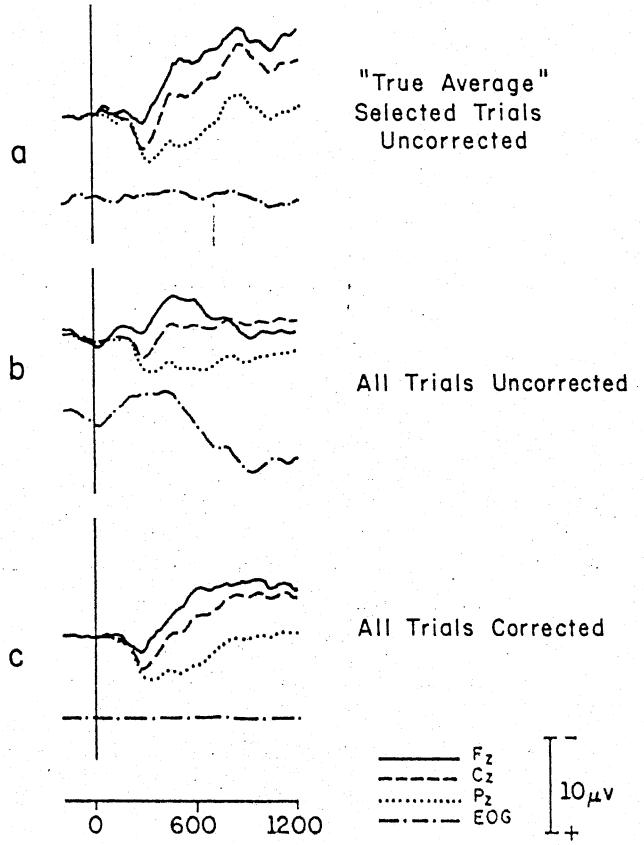

Fig. 5. 'True' ERP (a), and ERP derived from all the trials before (b) and after (c) correction by EMCP for subject 1. Associated EOG is also shown.

between trials following the application of EMCP. For subjects 4 and 5, about 1% of time points are associated with an increase in variance following correction. For subject 1 there is an increase following correction on 13% of time points. As a whole, about 99% of the time points show decrease in the variance between trials following correction.

Series 2

In the experiment which provided the data for the tests in series 1, the EOG electrodes were placed above and to the side of the eye. This oblique derivation is affected by horizontal, as well as by vertical, eye movements. However, we proposed (footnote 3) that because eye movements in this task were almost exclusively in the vertical direction, the EOG would reflect only vertical movements. Nevertheless, because of the possible contamination of the EOG record by horizontal movements in the first series of tests, we conducted a second series using data from an experiment in which a vertical placement was used for EOG. These tests involved data from a simple visual oddball task (Fabiani et al. 1982). The word 'count' was presented 100 times to each of 4 undergraduate female subjects, in each of 2 sessions, 1–2 weeks apart. On 20 occasions the word was larger than it was on the other 80 occasions. The subjects were instructed to count the number of 'rare' words (those with larger size). EEG was recorded at Fz, Cz and Pz, referred to linked

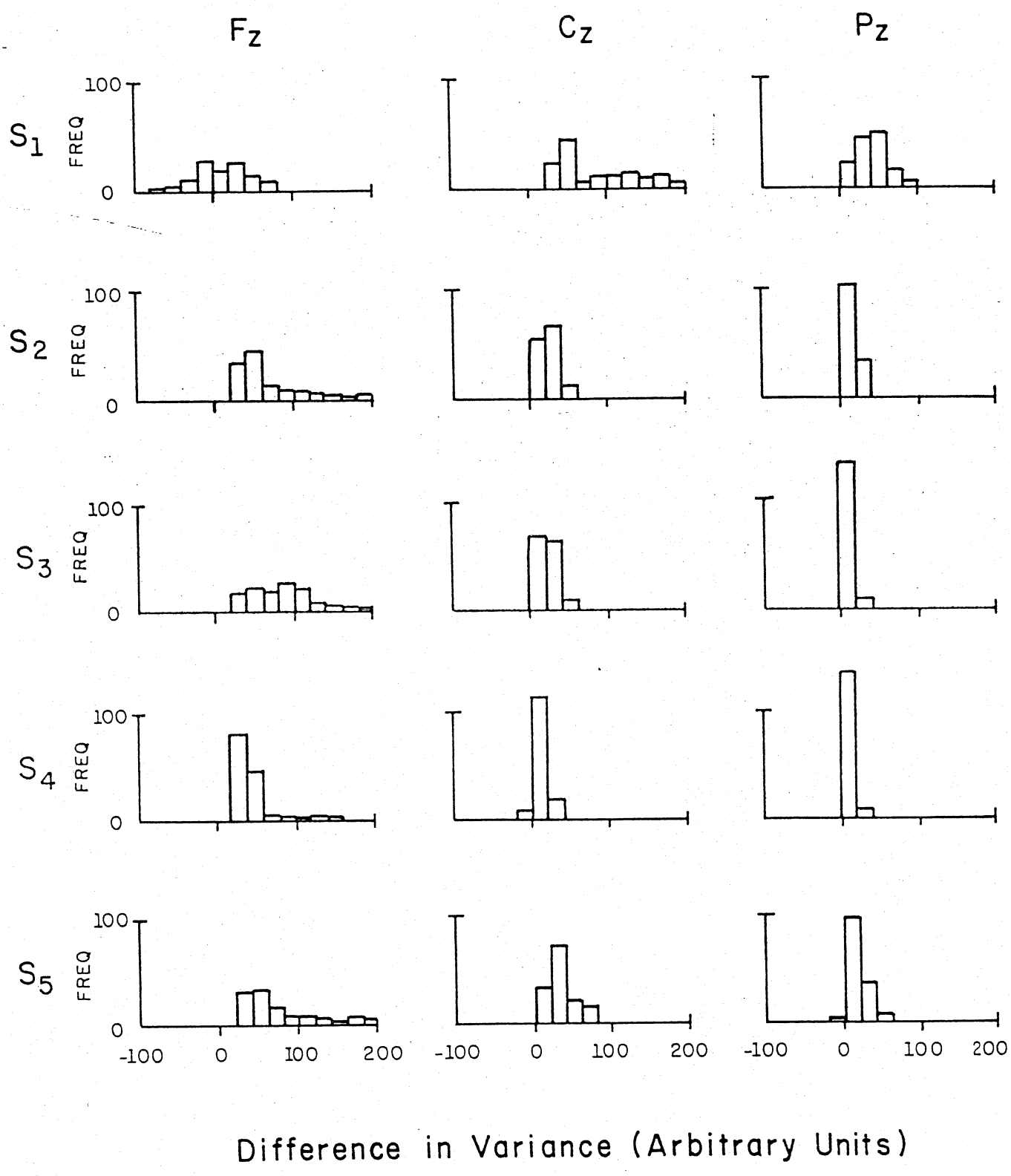

Fig. 6. Frequency distributions of the differences between the variance of corrected and uncorrected trials for each time point. Separate histograms are shown for each subject and electrode.

TABLE I

Deviations between ERPs based on large and small * EOG variance trials before and after correction.

Subject no.	Deviation before correction				% of trials with small EOG variance
	Fz	Cz	Pz	Tot	
1	81.91	48.02	37.81	59.01	25
2	51.75	41.26	28.79	41.67	77
3	20.00	14.24	12.53	15.91	52
4	43.31	33.69	24.00	34.58	56
5	37.07	25.13	18.69	28.02	50
	Deviation after correction				
	Fz	Cz	Pz	Tot	
1	28.84	19.89	16.84	22.44	
2	41.23	36.56	25.95	35.17	
3	14.21	11.86	10.68	12.34	
4	31.62	29.15	20.55	27.52	
5	21.24	23.49	16.79	20.70	

* Note: ERPs based on small EOG variance trials before correction correspond to the 'true' ERPs.

mastoids, and vertical EOG was recorded from electrodes placed above and below the right eye.

Data for each session were analyzed independently and averages were computed for 'rare' and 'frequent' stimuli and for each electrode separately. Three kinds of averages were computed:

(a) True ERP

Average of the records for trials during which EOG activity did not exceed a strict variance criterion.

(b) Raw ERP

Average of the records for all the trials, regardless of the occurrence of ocular artifacts.

(c) Corrected ERP

Average of the records for all the trials, regardless of the occurrence of ocular artifacts, but corrected according to EMCP.

The correction factors yielded by EMCP are presented in Table II. Separate correction factors are presented for each subject, session, scalp electrode and for blinks and saccades. Means and standard deviations for each scalp electrode and for blinks and saccades are also shown.

As can be seen in Table II, the correction factors for saccades are consistently larger than those for blinks, in accordance with the propagation factors reported in the literature (Corby and Kopell 1972). The correction factors both for blinks and saccades appear to be stable, although small differences between subjects can be observed. The correlations between the correction factors for the first and second sessions were 0.97 for blinks and

TABLE II

Correction factors from the visual oddball experiment.

Subject no.	Session no.	Blinks			Saccades		
		Fz	Cz	Pz	Fz	Cz	Pz
1	1	0.20	0.07	0.04	0.29	0.11	0.07
	2	0.19	0.07	0.04	0.23	0.10	0.05
2	1	0.13	0.08	0.05	0.18	0.10	0.06
	2	0.11	0.06	0.07	0.10	0.05	0.03
3	1	0.21	0.09	0.04	0.23	0.09	0.03
	2	0.19	0.07	0.02	0.23	0.10	0.04
4	1	0.16	0.05	0.03	0.20	0.07	0.03
	2	0.18	0.04	0.00	0.25	0.07	0.02
Mean		0.17	0.07	0.04	0.21	0.09	0.04
S.D.		0.03	0.01	0.02	0.05	0.02	0.02

0.91 for saccades (n = 12, $P < 0.01$). Note that the correction factor may change between experimental sessions if electrodes are not placed in exactly the same position because propagation characteristics will change with electrode position. This is particularly true if there are variations in the placement of the EOG electrodes. Furthermore, variation in impedances between sessions might also lead to instability in the estimate of correction factors. We also would not necessarily expect to find identical values for different subjects because of morphological differences.

Wave forms for the first subject are presented in Fig. 7. In this figure the averages from each session both for rare and frequent stimuli are presented. 'True,' 'raw' and 'corrected' ERPs for Fz, Cz and Pz are presented, together with the vertical EOG trace. This figure shows that a large ocular artifact is present at a latency of about 800 msec for the rare stimuli. It is particularly evident at the frontal electrode. Note that the artifact has been entirely corrected by EMCP.

To quantify the 'goodness' of the correction, deviation indices from the 'true ERPs' (defined as above) were computed for both the 'raw' and the 'corrected' ERPs, for each subject, session and electrode. The deviation indices are presented in Table III. As can be seen in this table, EMCP generally reduces the deviations from the 'true' ERPs. In a few cases the difference between the deviations for the 'corrected' and 'raw' ERPs is slightly negative. In this case the subjects made very few eye movements, and almost no trials were actually rejected in the computation of the 'true'

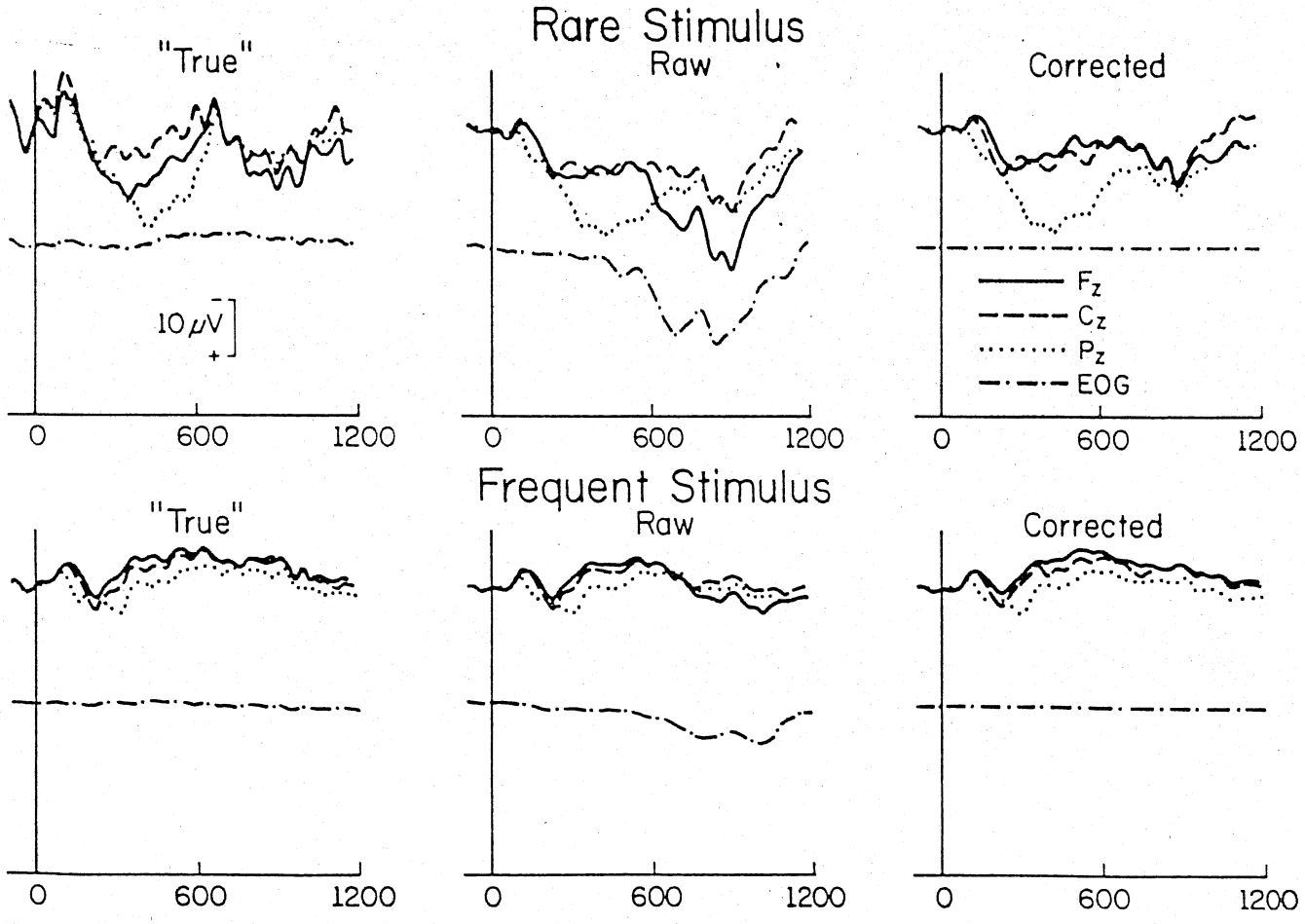

Fig. 7. 'True' ERPs, ERPs based on all trials, and corrected ERPs, for rare and frequent stimuli in the visual oddball experiment. Data for subject 1, session 1.

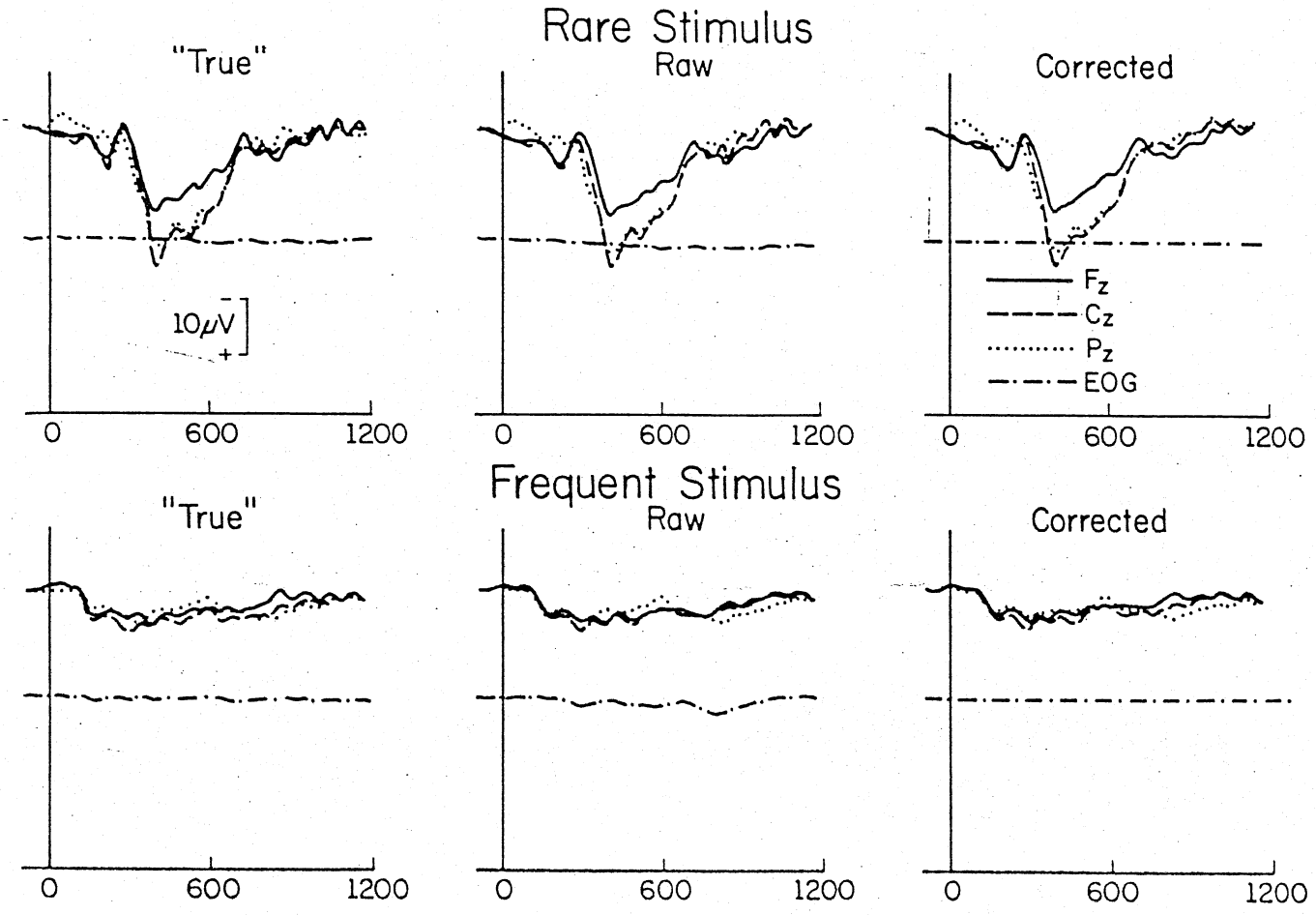

Fig. 8. 'True' ERPs, ERPs based on all trials, and corrected ERPs, for rare and frequent stimuli in the visual oddball experiment. Data for subject 4, session 1.

ERP. To illustrate this case, wave forms from the first session of the fourth subject are presented in Fig. 8.

Series 3

In the third series of tests, we assessed the constancy of the correction factor over different blocks of trials obtained from the same subject in the same experimental session as well as between sessions. The data used for this test series came from another experiment (Coles et al. in preparation). Four subjects performed a total of 2160 trials of a choice reaction time task in blocks of 108 trials over 3 or 4 experimental sessions. Each trial was initiated by depression of a foot pedal by the subject. One second later a visual display containing target letters S or H appeared on a projection tachistoscope and the subject made a left- or right-hand squeeze response. EOG and EEG, recorded from Fz, Cz and Pz, were obtained as in the previously described experiment. The epoch for analysis began 1280 msec before the foot press and extended through 1280 msec after the stimulus, making a total duration of 3560 msec.

For the purpose of analysis, each block of 108 trials was treated separately and individual correction factors for each electrode were computed. Fig. 9 shows the mean and inter-quartile range for correction factors for each subject and electrode separately over all trial blocks. Note that there is no overlap between electrodes, with correction factors for Fz > Cz > Pz. Note also that there is

TABLE III

Deviations from the 'true' ERP before and after correction.

Subject no.	Session no.	Deviation before correction			
		Fz	Cz	Pz	Total
1	1	69.29	38.02	33.48	49.56
	2	61.94	38.98	27.23	45.08
2	1	28.74	20.44	17.67	22.78
	2	6.04	5.61	4.51	5.43
3	1	30.19	23.42	15.12	23.73
	2	59.87	60.52	39.86	54.27
4	1	7.81	8.20	8.78	8.28
	2	24.66	19.90	17.44	20.88
		Deviation after correction			
		Fz	Cz	Pz	Total
1	1	31.72	33.75	24.93	30.37
	2	31.20	34.84	27.76	31.40
2	1	19.82	17.91	15.47	17.83
	2	5.63	5.20	4.38	5.10
3	1	21.16	24.50	12.85	20.11
	2	35.38	28.21	27.38	30.53
4	1	7.07	9.04	9.04	8.43
	2	15.27	19.35	17.27	17.38

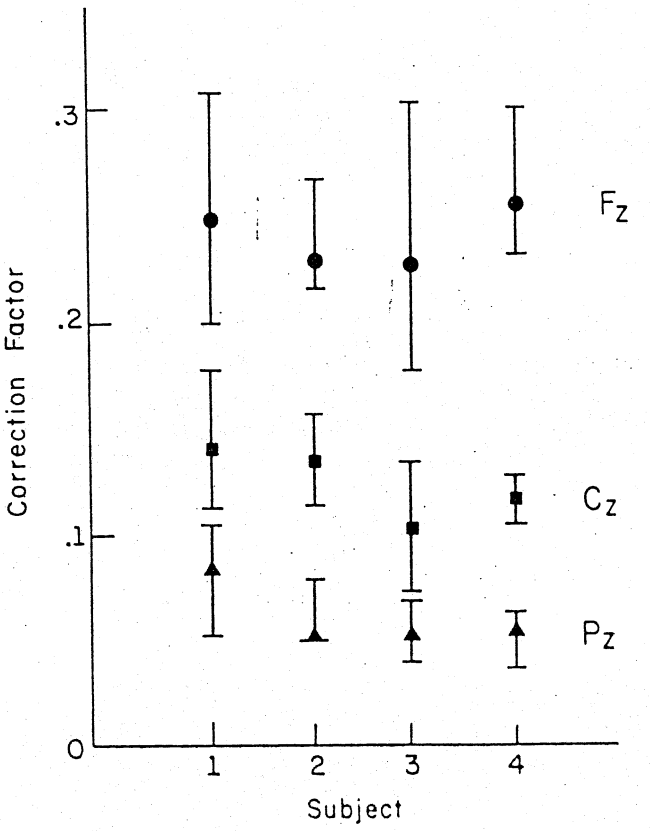

Fig. 9. Mean and interquartile range of the correction factors for 4 subjects for each electrode.

stability in correction factors between subjects within electrodes. The within-subject variability seen in Fig. 9 is mostly due to between-session variability. This is illustrated in Fig. 10 where the data for subject 3 are shown for each session (day) separately. Note that there is some variability across days, but that the variability within a session is smaller than that between sessions. As noted above, the variability across days may be due to slight variations in electrode placement and/or impedance between sessions. This variability points to the importance of computing separate correction factors at least for each experimental session.

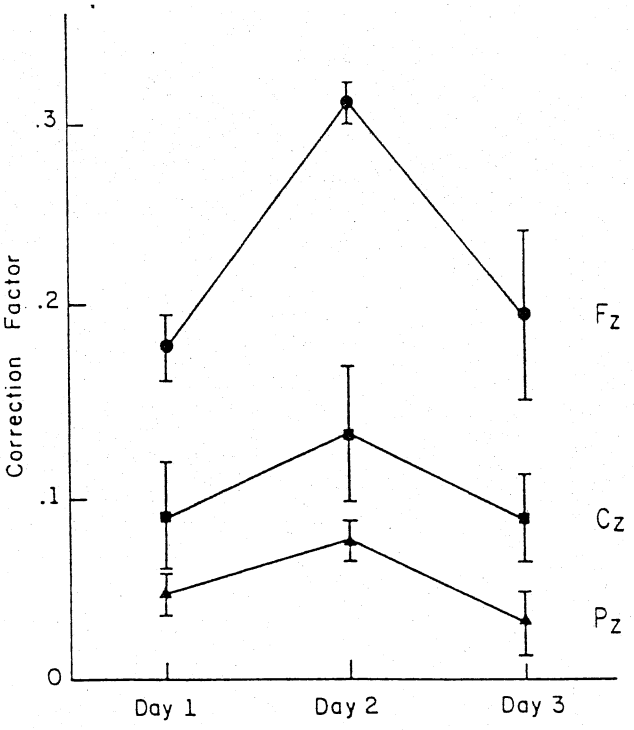

Fig. 10. Subject 3: mean and standard deviation of the correction factors for each electrode as a function of session (day).

Discussion

The results of the first and second series of tests reveal that trials corrected by EMCP yield an ERP which corresponds more closely to the 'true' ERP

than ERPs derived from 'randomly' corrected or uncorrected trials. This is true for trials selected on the basis of large eye movement or blink contamination and for unselected trials. For both trial samples, EMCP reduces deviation from the 'true' ERP, but some deviation remains. This residual deviation is not surprising. As we have noted, our estimate of the 'true' ERP may not correspond to the 'true' ERP. First, this estimate may be computed on too small a number of trials to eliminate random noise from the average. Secondly, any traditional rejection criterion, such as that used here, can permit trials associated with small, and perhaps consistent, eye movement artifact to be included in the average. A rejection criterion of zero variance in the EOG channel is clearly unrealistic. Thus, some contamination may be present even in an average based on trials selected according to a strict EOG variance criterion.

A second reason for the persistence of some deviation from the 'true' ERP even after correction is, of course, that the correction factors are themselves estimates, although based, in our procedure, on a large sample. Noise in the EOG channel could significantly reduce the magnitude of the correction factor.

Other tests of validity also support our claim that EMCP is a valid procedure. The difference between ERPs derived from trials with different degrees of artifact decreases following correction. Furthermore, variance between individual time points across trials decreases following application of EMCP. In both cases, the procedure reduces, but does not eliminate, variance between ERPs or among single trials. Again, this is not surprising. Eye movement or eye blink artifacts are not the only source of variance contributing to differences between ERPs or among single trials.

The validity of EMCP is also supported by the data obtained from the reliability tests (second and third series). Because of the propagation characteristics of the head, we would expect that ERPs obtained from frontal electrode sites would be more contaminated than those obtained from more posterior derivations (cf., Overton and Shagass 1969). Figs. 9 and 10 and Table II indicate clearly that the magnitude of the correction factor decreases from frontal to parietal electrodes. The second series of tests also show that the correction factors are consistently larger for saccades than for blinks (Corby and Kopell 1972).

The data from the third series reveal that values of the correction factor are stable within subjects and sessions, although there is some variability between sessions.

Thus, we believe that EMCP is both a valid and reliable procedure for dealing with the problem of eye movement and eye blink artifact. Furthermore, the procedure has clear advantages. First, it distinguishes between blink and eye movement artifact. Second, by deleting stimulus-related EOG and EEG activity before computing the correction factors it provides corrections that are insensitive to stimulus-locked activity at the EOG electrodes. Third, *all* trials are retained for use in subsequent analyses. Fourth, it is a general procedure which does not require special data to be collected. In this respect, it also reduces the problem of a possible difference between voluntary and involuntary movements and blinks. Fifth, the subjects need not be instructed to 'control' or minimize their eye movements. As we have noted, such instructions may impose an unwanted secondary task. Sixth, the estimate of artifact is based on a large sample (trials × time points) rather than a few points obtained from a few prescribed eye movements (cf., Hillyard and Galambos 1970; Girton and Kamiya 1973).

The EMCP has been implemented in a Fortran program on the Harris/7 computer at the Cognitive Psychophysiology Laboratory. For the data used in the validity tests described above (Fabiani et al. 1982; Donchin et al. 1983), computation of correction factors for each subject and subsequent correction takes approximately 5 min of CPU time. These data consisted of 100 trials, 140 points/trial, for 3 EEG channels, 1 EOG channel and 1 event channel. For comparison, we note that simple averaging of the same data set takes 1.5 min of CPU time, while latency adjustment using the Woody technique (3 iterations) and step-wise discriminant analysis take 4 min and 2.5 min, respectively. Thus, the EMCP technique does consume considerable computer resources. However, the gains in precision and data preservation, in our opinion, more than justify the costs.

The procedure could be made both more simple or more complicated. If blinks are not a particular problem, then trials on which blinks occur could be eliminated and EMCP performed only to correct for movements. This would result in a 50–75% reduction in computer time. On the other hand, the procedure could be more complex if up and down movements were considered individually. This would also require monopolar recordings of vertical EOG.

The advantages of the procedure listed above suggest that EMCP could be fruitfully used in any situation in which ERPs are recorded. However, there are certain situations where it seems particularly appropriate. These situations include those where eye movements are part of the task or are characteristic of the population being studied. When the procedure is used, our tests suggest that individual correction factors should be computed for each subject, electrode and experimental session. Where ERPs from midline sites are of interest, only vertical artifacts need to be corrected. However, if lateral derivations are used, then both vertical and horizontal EOG should be recorded and used to compute separate correction factors for blinks and vertical and horizontal eye movements.

Finally, we should note that the procedure has been used by us in the analysis of the data from Donchin et al. (1983, see above). The benefits gained were very apparent. First, *all* trials could be used in analysis — even though, for one condition, only 5% of trials were free of eye movements and blinks. Second, condition differences which were evident, but not reliable, in analyses of data selected by traditional rejection procedures, became reliable when the analyses were based on corrected trials.

Summary

A new off-line procedure for dealing with ocular artifacts in ERP recording is described. The procedure (EMCP) uses EOG and EEG records for individual trials in an experimental session to estimate a propagation factor which describes the relationship between the EOG and EEG traces. The propagation factor is computed after stimulus-linked variability in both traces has been removed. Different propagation factors are computed for blinks and eye movements.

Tests are presented which demonstrate the validity and reliability of the procedure. ERPs derived from trials corrected by EMCP are more similar to a 'true' ERP than are ERPs derived from either uncorrected or randomly corrected trials. The procedure also reduces the difference between ERPs which are based on trials with different degrees of EOG variance. Furthermore, variability at each time point, across trials, is reduced following correction.

The propagation factor decreases from frontal to parietal electrodes, and is larger for saccades than blinks. It is more consistent within experimental sessions than between sessions.

The major advantage of the procedure is that it permits retention of all trials in an ERP experiment, irrespective of ocular artifact. Thus, studies of populations characterized by a high degree of artifact, and those requiring eye movements as part of the experimental task, are made possible. Furthermore, there is no need to require subjects to restrict eye movement activity. In comparison to procedures suggested by others, EMCP also has the advantage that separate correction factors are computed for blinks and movements and that these factors are based on data from the experimental session itself rather than from a separate calibration session.

Résumé

Nouvelle méthode d'élimination off-line des artéfacts oculaires

Une nouvelle technique 'off-line' pour éliminer les artéfacts oculaires lors de l'enregistrement des PER est décrite. La technique 'EMCP' utilise les enregistrements de l'EOG et de l'EEG au cours d'essais individuels d'une session expérimentale, pour déterminer un facteur de propagation décrivant la relation entre tracés EOG et EEG. Le facteur de propagation est calculé après que la variabilité liée au stimulus est été éliminée sur les

deux tracés. Différents facteurs de propagation sont calculés pour les clignements de paupières et les mouvements oculaires.

Des tests démontrant la validité et la fiabilité de la technique sont présentés. Les PER obtenus à partir d'essais corrigés par la EMCP ressemble davantage à un 'vrai' PER qu'un PER obtenu à partir d'essais, soit non corrigés, soit corrigés au hasard. La méthode réduit également la différence entre PER recueillis lors d'essais avec différents degré de variance de l'EOG. De plus la variabilité à chaque instant entre les essais est réduite après correction.

Le facteur de propagation décroît des électrodes frontales aux électrodes pariétales, il est plus grand pour des saccades que pour des clignements. Il est plus stable à l'intérieur d'une session que d'une session à l'autre.

L'avantage majeur de la méthode est de pouvoir retenir tous les essais dans une expérience de PER, quels que soient les artéfacts oculaires. Ainsi les études de populations caractérisées par un grand nombre d'artéfacts et celles nécessitant des mouvements oculaires comme partie intégrante de l'expérience, sont devenues possibles. De plus, il n'est pas nécessaire de demander au sujet de limiter son activité motrice oculaire.

Par rapport à des techniques proposées par d'autres l'EMCP a aussi l'avantage d'établir des facteurs de correction séparés pour les clignements de paupières et les mouvements oculaires. Ces facteurs sont basés sur des données recueillies lors de la séance expérimentale elle-même, et non à partir d'une séance d'étalonnage séparée.

Earle Heffley, Amir Mane and Noel Marshall made helpful comments on earlier versions of the manuscript. In developing the procedure for distinguishing between eye movements and blinks, we relied on unpublished work by Ron Clapman in our laboratory.

References

Barry, W. and Jones, G.M. Influence of eyelid movement upon electro-oculographic recording of vertical eye movements. Aerospace Med., 1965, 36: 855–858.

Carpenter, A. The rate of blinking during prolonged visual search. J. exp. Psychol., 1948, 38: 587–591.

Corby, J.C. and Kopell, B.S. Differential contributions of blinks and vertical eye movements as artifacts in EEG recording. Psychophysiology, 1972, 9: 640–644.

Donchin, E., Callaway, E., Cooper, R., Desmedt, J.E., Goff, W.R., Hillyard, S.A. and Sutton, S. Publication criteria for studies of evoked potentials (EP) in man. Report of the Methodology Committee. In: J.E. Desmedt (Ed.), Attention, Voluntary Contraction and Event-related Cerebral Potentials. Prog. clin. Neurophysiology, Vol. 1. Karger, Basel, 1977: 1–11.

Donchin, E., Coles, M.G.H. and Gratton, G. Cognitive psychophysiology and preparatory processes: a case study. In: S. Kornblum and J. Requin (Eds.), Preparatory States and Processes. Erlbaum, Hillsdale, N.J., 1983: in press.

Drew, G.C. Variations in reflex blink-rate during visual motor tasks. Quart. J. exp. Psychol., 1951, 3: 75–88.

Fabiani, M., Karis, D. and Donchin, E. P300 and memory: individual differences in the Von Restorff effect. Psychophysiology, 1982, 19: 558 (abstract).

Girton, D.G. and Kamiya, J. Simple on-line technique for removing eye movement artifacts from the EEG. Electroenceph. clin. Neurophysiol., 1973, 34: 212–216.

Hillyard, S.A. and Galambos, R. Eye movement artifact in the CNV. Electroenceph. clin. Neurophysiol., 1970, 28: 173–182.

Mowrer, O.H., Ruch, T.C. and Miller, N.E. The corneo-retinal potential differences as the basis of the galvanic method of recording eye movements. Amer. J. Physiol., 1935, 114: 423–428.

Overton, D.A. and Shagass, C. Distribution of eye movement and eyeblink potentials over the scalp. Electroenceph. clin. Neurophysiol., 1969, 27: 546 (abstr.).

Records, R.E. Eyebrows and eyelids. In: R.E. Records (Ed.), Physiology of the Human Eye and Visual System. Harper and Row, Hagerstown, 1979: 1–24.

Verleger, R., Gasser, T. and Möcks, J. Correction of EOG artifacts in ERP of the EEG: aspects of reliability and validity. Psychophysiology, 1982, 19: 472–480.

Appendix

Derivation of the EMCP procedure

Between any given electrode pair, on any trial, we record for each of the time points a voltage V_t, which can be expressed as

$$V_t = ERP_t + EMA_t + N_t \quad (1)$$

where ERP_t = the 'true' value of the ERP at time point t; EMA_t = the contribution to V_t of signals generated at the eyeball; N_t = the 'noise' which includes all electrical activity not time locked to the eliciting event and not associated with oculographic signals. It is assumed that the expected

value of N_t is 0. Note that $\bar{V}_t = E(V_t) = ERP_t + \overline{EMA}_t$, where EMA_t is the ensemble average of the individual EMAs. As there is no reason to believe that $E(EMA_t) = 0$, the EMA_t is an artifact that must be removed from V_t or $E(V_t)$.

We assume that EMA_t is a scaled value of the signal recorded at the eyeball by the oculographic electrode pair. Thus,

$$EMA_t = K(EOG_t) \tag{2}$$

where EOG_t is the value recorded at the ocular electrodes at time point t; K is the scaling constant that determines the contribution of the EOG signal at V_t.

It is the goal of our procedure to estimate K. This value can then be used to 'correct' V_t to remove $\overline{EMA}_t$ and to allow proper interpretation of the data. The estimation problem is of course complicated by the fact that EMA_t is unknown. We solve this problem in the following way.

We assume that the EOG is composed of two components as follows

$$EOG_t = EMR_t + EMN_t \tag{3}$$

where EMR_t = the voltage generated by an eye movement triggered, and time-locked, to the eliciting event; EMN_t = voltage at time point t that is not time-locked to the eliciting event. It is assumed that $E(EMN_t) = 0$.

We can now express V_t as follows:

$$V_t = ERP_t + K(EMR_t + EMN_t) + N_t \tag{4}$$

Therefore

$$E(V_t) = ERP_t + K(EMR_t) \tag{5}$$

Of course, $\bar{V}_t$, the ensemble average of the V_t over the n trials, is an estimate of $E(V_t)$.

We can now obtain the following relations. From (5) and (4), we get

$$V_t - \bar{V}_t = K(EMN_t) + N_t \tag{6}$$

from (6) and (3) we get

$$V_t - \bar{V}_t = K(EOG_t - EMR_t) + N_t \tag{7}$$

As EMR_t can be estimated from the ensemble average of the EOG_t, by the same reasoning that estimates the ERP_t by $\bar{V}_t$, we get

$$V_t - \bar{V}_t = K(EOG_t - \overline{EOG}_t) + N_t \tag{8}$$

Note that all elements in equation (8), with the exception of K and N_t, are given by the data. However, as $E(N_t) = 0$, we can estimate K using standard least-squares techniques. The correction factor K is therefore obtained by

(1) Computing $\bar{V}_t$ and $\overline{EOG}_t$, for all p time points, over the trials.

(2) Computing, for each trial, $(V_t - \bar{V}_t)$ and $(EOG - \overline{EOG}_t)$.

(3) Solving, using least-squares regression, K in the equation, $(V - \bar{V}_t) = K(EOG - \overline{EOG}_t)$, using all the pairs $[(V_t - \bar{V}_t), (EOG - \overline{EOG}_t)]$ generated by all trials and all time points.

The equation is solved separately for time points affected by blinks and those affected by eye movements. A blink is assumed to have occurred at time point t, whenever the EOG signals exceed a preset criterion within a 20 msec interval bracketing time point t. Thus, if $(EOG_t - EOG_{t-10}) + (EOG_t - EOG_{t+10}) \geq$ criterion, a blink is assumed to have occurred.