



J Neurophysiol. 2011 December; 106(6): 3216–3229.

PMCID: PMC3234084

Published online 2011 August 31. doi: [10.1152/jn.00220.2011](https://doi.org/10.1152/jn.00220.2011)

Taking into account latency, amplitude, and morphology: improved estimation of single-trial ERPs by wavelet filtering and multiple linear regression

[L. Hu](#),^{1,2,3} [M. Liang](#),² [A. Mouraux](#),⁴ [R. G. Wise](#),⁵ [Y. Hu](#),³ and [G. D. Iannetti](#)^{1,2}✉

¹Key Laboratory of Cognition and Personality (Ministry of Education) and School of Psychology, Southwest University, Chongqing, China;

²Department of Neuroscience, Physiology and Pharmacology, University College London, United Kingdom;

³Department of Orthopaedics and Traumatology, The University of Hong Kong, Hong Kong, China;

⁴Institute of Neurosciences, Université catholique de Louvain, Brussels, Belgium; and

⁵Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, Cardiff, United Kingdom

✉Corresponding author.

Address for reprint requests and other correspondence: G. D. Iannetti, Dept. of Neuroscience, Physiology and Pharmacology, Univ. College London, Medical Sciences Bldg., Gower St., WC1E 6BT London (e-mail: g.iannetti@ucl.ac.uk).

Received March 14, 2011; Accepted August 30, 2011.

Copyright © 2011 the American Physiological Society

Abstract

Across-trial averaging is a widely used approach to enhance the signal-to-noise ratio (SNR) of event-related potentials (ERPs). However, across-trial variability of ERP latency and amplitude may contain physiologically relevant information that is lost by across-trial averaging. Hence, we aimed to develop a novel method that uses 1) wavelet filtering (WF) to enhance the SNR of ERPs and 2) a multiple linear regression with a dispersion term (MLR_d) that takes into account shape distortions to estimate the single-trial latency and amplitude of ERP peaks. Using simulated ERP data sets containing different levels of noise, we provide evidence that, compared with other approaches, the proposed WF+MLR_d method yields the most accurate estimate of single-trial ERP features. When applied to a real laser-evoked potential data set, the WF+MLR_d approach provides reliable estimation of single-trial latency, amplitude, and morphology of ERPs and thereby allows performing meaningful correlations at single-trial level. We obtained three main findings. First, WF significantly enhances the SNR of single-trial ERPs. Second, MLR_d effectively captures and measures the variability in the morphology of single-trial ERPs, thus providing an accurate and unbiased estimate of their peak latency and amplitude. Third, intensity of pain perception significantly correlates with the single-trial estimates of N2 and P2 amplitude. These results indicate that WF+MLR_d can be used to explore the dynamics between different ERP features, behavioral variables, and other neuroimaging measures of brain activity, thus providing new insights into the functional significance of the different brain processes underlying the brain responses to sensory stimuli.

Keywords: event-related potentials, laser-evoked potentials, single-trial analysis, multiple linear regression with dispersion term

EVENT-RELATED POTENTIALS (ERPs) are transient changes in the ongoing electroencephalogram (EEG) elicited by sensory, motor, or cognitive events. ERPs largely reflect synchronous changes of slow postsynaptic potentials occurring within a large number of similarly oriented cortical pyramidal neurons ([Nunez and Srinivasan 2006](#)). Because the magnitude of ERPs is often several factors smaller than the

magnitude of the background EEG ([Hu et al. 2010](#)), their identification relies on signal processing methods for enhancing the signal-to-noise ratio (SNR). The most widely used approach to enhance SNR is across-trial averaging in the time domain ([Dawson 1951, 1954](#)). The validity of this across-trial averaging approach relies on two basic assumptions ([Spencer 2005](#)): 1) that ERPs are stationary (i.e., the latency and morphology are invariant across trials) and 2) that ERPs are independent of the background EEG activity. However, ERPs comprise multiple waves whose latency and amplitude can greatly and independently vary from trial to trial ([Spencer 2005](#)), and there is convincing evidence that background EEG is often correlated with the ERP response (e.g., because of phase resetting of ongoing EEG oscillations; [Makeig et al. 2002](#); [Sauseng et al. 2007](#)). For these reasons, the cost of the across-trial averaging approach is that all the information concerning across-trial variability of ERP latency and amplitude is lost ([Mouraux and Iannetti 2008](#)). Thus across-trial averaging can heavily bias the representation of cortical activity elicited by a sensory stimulus, because across-trial variability often contains physiologically relevant information, related, for example, to changes in stimulus parameters (duration, intensity, and location) and fluctuations in vigilance, expectation, attentional focus, or task strategy ([Lee et al. 2009](#); [Legrain et al. 2002, 2003](#)). Therefore, the ability to obtain a reliable single-trial estimate of ERP latency and amplitude is highly desirable, because it would allow exploration of the single-trial dynamics between ERP measures and behavioral variables (e.g., intensity of perception, reaction time) ([Iannetti et al. 2005b](#)), as well as measurements of brain activity obtained using different neuroimaging techniques (e.g., functional MRI) ([Mayhew et al. 2010](#)).

Taking into account across-trial variability is of particular importance for long-latency ERPs, such as the ERPs elicited by nociceptive laser stimuli (laser-evoked potentials, LEPs). LEPs are considered the best tool for assessing function of nociceptive pathways in physiological and clinical studies ([Bromm and Treede 1991](#); [Cruccu et al. 2008](#); [Iannetti et al. 2001](#)), because laser heat pulses excite selectively A δ and C fiber-free nerve endings in the superficial skin layers (i.e., without coactivating A β mechanoreceptors) ([Bromm and Treede 1984](#); [Carmon et al. 1976](#)). LEPs have been shown to be related to the activation of slow-conducting type-II A δ mechano-heat nociceptors ([Treede et al. 1998](#)) and spinothalamic neurons located in the anterolateral quadrant of the spinal cord ([Iannetti et al. 2003](#); [Treede et al. 2003](#)). LEPs comprise a number of waves that are time locked to the onset of the stimulus. The largest response is a negative-positive vertex potential (N2 and P2 waves, peaking at ~200 and 350 ms when stimulating the hand dorsum) ([Bromm and Treede 1984](#)). The latency, amplitude, and morphology of the N2 and P2 waves of LEPs exhibit especially high across-trial variability ([Iannetti et al. 2005b](#); [Purves and Boyd 1993](#)), most probably due to a unique combination of peripheral (e.g., time-dependent fluctuations in baseline skin temperature, variability in the number of activated nociceptive fibers, variability in conduction velocity resulting in differences in the spatial summation at central synapses) and cognitive factors (e.g., fluctuations in vigilance, attentional focus and task strategy) ([Lee et al. 2009](#); [Legrain et al. 2002, 2003](#)). For all these reasons, a reliable single-trial estimate of LEP latency and amplitude is highly desirable, and LEPs represent an ideal model to develop novel approaches to analyze ERPs at a single-trial level.

In a previous study, we described a method based on multiple linear regression (MLR) to estimate automatically latency and amplitude of single-trial ERPs ([Mayhew et al. 2006](#)). This method has been successfully applied to the single-trial detection of the N2 and P2 waves of LEPs ([Mayhew et al. 2006](#)) and the N1 and P2 waves of auditory evoked potentials ([Mayhew et al. 2010](#)). More recently, we combined time-frequency wavelet filtering and MLR to enhance the SNR and thus achieve a more robust single-trial estimate of ERP components, making it possible to explore the single-trial dynamics of components having a very small SNR, such as the N1 wave of LEPs ([Hu et al. 2010](#)). Such an MLR approach is a procedure commonly used to analyze functional MRI data ([Friston et al. 1998](#)), where not only the canonical hemodynamic response function (analogous to the average ERP in this case) but also

its temporal derivative (to account for the temporal variability of the hemodynamic response) is fitted to the data in a general linear model (GLM) framework. Thus the inclusion of the temporal derivative in the regressors allows capturing of not only the amplitude but also the latency jitter of the single-trial ERP responses. However, besides latency and amplitude, another factor that varies from trial to trial is the morphology of the ERP waveform (Casarotto et al. 2005; Jung et al. 2001; Mouraux and Iannetti 2008). In addition, the significant latency jitter may lead to an important distortion of the averaged ERPs, which thus will not constitute a faithful representation of the single-trial ERP response. This is relevant in clinical ERP studies, because many abnormal conditions (e.g., optic neuritis in multiple sclerosis) are characterized by the so-called “desynchronized” ERPs, i.e., ERPs with reduced amplitude, longer latency, and, importantly, increased width of their waves (Orssaud 2003; Pelosi et al. 1997). These variations in ERP morphology (i.e., wave width) could be an important parameter to be considered and quantified when using ERPs as a diagnostic tool in clinical studies. Therefore, to obtain a more accurate single-trial estimate of ERPs, it would be desirable to take into account not only the variability of the latency and amplitude (as in Mayhew et al. 2006 and Hu et al. 2010) but also the variability in morphology of the ERP waveform.

To address this need and thereby provide an improved single-trial estimate of ERP components, we propose a novel single-trial detection method based on a multiple linear regression that includes a dispersion term (MLR_d). In this method, a data-driven nonparametric approach based on principal component analysis (PCA) was used to identify the principal components that capture the variance of both latency and morphology of single-trial ERP waveforms. These principal components were then used to generate a set of regressors that were fitted to the single-trial ERP waveforms. In addition, time-frequency WF, which has been proved effective to enhance the SNR of ERPs in both single-trial and average waveforms (Hu et al. 2010), was used before the application of MLR_d.

We compared the performance of this novel method (WF+MLR_d) with previous versions of the MLR that do not take into account the variability in morphology to estimate single-trial ERP information (MLR, WF+MLR) on both simulated ERP data sets with variable levels of noise and a real LEP data set. Finally, we used this novel method to explore the relationship between the intensity of pain perception and the latency and amplitude of the N2 and P2 LEP waves at the level of single trials.

MATERIALS AND METHODS

Single-Trial Analysis Approaches

The wavelet filtering (WF) method is used for enhancing the SNR of ERPs in both single trials and averages (Hu et al. 2010). The methods based on multiple linear regression (MLR) and multiple linear regression with dispersion term (MLR_d) for estimating automatically the latency and amplitude of ERPs in single trials are summarized in Fig. 1. In this study we assessed the performance of MLR and MLR_d, with and without preceding WF, on simulated data sets. The method with best performance was applied to extract single-trial values from a real LEP data set. All these methods have been developed into user-friendly software running under the MATLAB environment and can be freely downloaded from the IannettiLab website (<http://iannettilab.webnode.com>).

Multiple linear regression. To estimate the latency and amplitude of single-trial ERP peaks in an unbiased fashion, we applied a linear regression method similar to the one we originally described previously (Mayhew et al. 2006) to both WF and non-WF filtered ERP data within the 0- to 500-ms poststimulus time interval (Fig. 1, top). The MLR approach takes into account the variability of both latency and amplitude of ERP. This variability can be described as follows:

$$f(t) = k_N y_N(t + a_N) + k_P y_P(t + a_P)$$

(1)

where $f(t)$ is a single-trial ERP waveform that varies as a function of time t . $f(t)$ can be modeled by the sum of the varied version of N2 wave [$k_N y_N(t + a_N)$] and P2 wave [$k_P y_P(t + a_P)$]. k_N and k_P are the weighted constants of N2 wave and P2 wave, whereas a_N and a_P are the latency shift values of the N2 wave and P2 wave, respectively. Since the N2 and P2 peaks of the LEPs reflect the activity of different neural generators ([Garcia-Larrea et al. 2003](#)), and their amplitude can be differentially modulated by several experimental factors (e.g., spatial attention and probability of perception) ([Lee et al. 2009](#); [Legrain et al. 2002](#)), we modeled the N2 and P2 waves separately, thus avoiding the assumption that all generators contributing to the LEP responses covary linearly ([Mayhew et al. 2006](#)).

By using the Taylor expansion, the MLR model can be written as

$$f(t) \approx k_N y_N(t) + a_N k_N y'_N(t) + k_P y_P(t) + a_P k_P y'_P(t) \quad (2)$$

where $y_N(t)$ and $y_P(t)$ are the averages of N2 and P2 waves and $y'_N(t)$ and $y'_P(t)$ are the temporal derivatives of N2 and P2 waves, respectively. Thus the single-trial ERP waveform is approximated using the sum of the weighted averages of the N2 and P2 waves and their respective temporal derivatives.

Based on the fitted waveform, latency and amplitude of the ERP response in each single trial were measured by finding the most negative peak if $k_N > 0$ (positive fit) or the most positive peak if $k_N < 0$ (negative fit) for the N2 wave, and the most positive peak if $k_P > 0$ (positive fit) or the most negative peak if $k_P < 0$ (negative fit) for the P2 wave, within a 100-ms time window centered on the latency of the N2 and P2 peak identified in the average waveform of each subject. Single-trial latencies were finally calculated from the latencies of the corresponding amplitude peaks.

Multiple linear regression with dispersion term. To obtain a more accurate single-trial estimate, it is desirable to take into account not only the variability of the latency and amplitude but also the variability in morphology of the ERP waveform. For this reason, we included a dispersion coefficient (representing the variability of waveform morphology, see [Fig. 1, bottom](#)) into the MLR model, thus obtaining a dispersed version of it (MLR_d):

$$f(t) = k_N y_N(s_N t + a_N) + k_P y_P(s_P t + a_P) \quad (3)$$

where s_N and s_P are the time dispersion coefficients that determine the compression ratios of the width of N2 and P2 waves of single-trial ERP compared with those of the average ERP, respectively.

However, it is difficult to estimate these parameters (i.e., s , a , and k) because the real function expressing single-trial LEP waveforms is unknown. For this reason, in the present study we employed a data-driven method, PCA, to define a basis set to fit the N2 and P2 ERP waves. This method is similar to the procedure commonly used to model the hemodynamic response in some blood oxygen level-dependent functional MRI (BOLD fMRI) studies ([Friman et al. 2003](#); [Hosseini-Zadeh et al. 2003](#); [Woolrich et al. 2004](#)). PCA is a mathematical approach that can transform the ERP data into several uncorrelated variables (PCs) ([Jolliffe 2002](#)). Three PCs for each ERP wave, mainly representing 1) the average ERP wave across trials, 2) the variability of the latency, and 3) morphology of single-trial ERPs, respectively,¹ are identified and then used to model the N2 and P2 waves. Therefore, this method includes the variability of latency and morphology of each ERP wave as regressors in the linear model.

The whole MLR_d procedure consists of the five steps outlined below ([Fig. 1, bottom](#)).

GENERATION OF THE VARIABILITY MATRICES. A large number of plausible responses were generated by simultaneously shifting (from -50 to +50 ms around the peak latency, in steps of 5 ms, for a total of 21 possible latency shifts) and compressing [from 1 to 2 (this value represents the compression ratio between the width of average waveform and the width of single-trial responses), centered at the peak latency, in steps of 0.05, for a total of 21 width changes] the average response of each subject in an enumerative fashion, thus creating 441 (i.e., 21×21) possible ERP responses in these set ranges. All these plausible responses were arranged into a matrix that we call the “variability matrix” (1 matrix for each ERP peak; [Fig. 1, step 1](#)). The sorting order of trials in the variability matrices is of no importance, because the order of the trials does not affect the PCA output. Note that the variability of ERP amplitude will be captured by the coefficients weighting the basis sets, and thus it is not included in the variability matrix.

PCA SEPARATION. The variability matrices were fed into PCA to obtain the PCs representing the linear subspace for the variability matrices ([Fig. 1, step 2](#)). Because the first few PCs capture most of the data set variance ([Hossein-Zadeh et al. 2003](#)), and the variability matrix was generated by varying latency and morphology of the average ERP waveform, it is expected that the first few PCs will represent the average ERP wave and the variations of its latency and morphology.

BASIS SET DEFINITION. Because the first three PCs always captured the properties of the average ERP wave and the variations of its latency and morphology, these were selected to constitute the basis sets for each ERP wave ([Fig. 1, step 3](#)). Note that basis sets with more regressors (PCs) would potentially provide not only a better fit of single-trial ERP waves but also a better fit of the noise in each single-trial waveform. In contrast, basis sets with fewer regressors (PCs) would potentially provide a worse fit of single-trial ERP waves but also would be less prone to fit the noise. In this study, the first three regressors (PCs) were selected because they mainly represent different physiological features of each ERP: its waveform, temporal jitter, and variability in morphology, respectively.

SINGLE-TRIAL FITTING. The obtained basis sets were regressed against the corresponding time window of each single ERP trial. The coefficient of each of the three regressors was estimated using the least-squares approach, as described by [Mayhew et al. \(2006\)](#) and [Hu et al. \(2010\)](#). By multiplying these estimated coefficients with the corresponding regressors, the fitted single-trial waveforms were reconstructed ([Fig. 1, step 4](#)).

SINGLE-TRIAL LATENCY AND AMPLITUDE MEASUREMENT. Finally, latency and amplitude of the ERP response in each single trial were measured by finding the most negative peak if $k_N > 0$ (positive fit) or the most positive peak if $k_N < 0$ (negative fit) for N2 wave, and the most positive peak if $k_P > 0$ (positive fit) or the most negative peak if $k_P < 0$ (negative fit) for P2 wave, within a 100-ms time window centered on the latency of the N2 and P2 peak in the average wave form of each subject. Single-trial latencies were calculated from the latencies of the corresponding amplitude peaks.

Single-Trial Performance on Simulated Data Set

The performance of the novel method described in the present article (i.e., combining WF with MLR_d) was compared with that achieved by using other methods available to estimate single-trial ERP information (i.e., MLR, MLR_d, and WF+MLR). This comparison was performed on simulated data. These simulated data were based on group-level ERP waveforms obtained by averaging EEG data from 12 subjects ([Fig. 2](#)). This grand average represents a typical ERP elicited by nociceptive laser stimulation, recorded at channel Cz and consisting of a biphasic negative-positive (N2-P2) response (N2 peak: 207 ms, -22.4 μ V; P2 peak: 364 ms, 12.9 μ V). The epoch length was 1,500 ms (baseline corrected using the prestimulus interval from -500 to 0 ms), and the sampling rate was 1,024 Hz.

Simulated data generation

The procedure to generate the simulated ERP data, which is summarized in [Fig. 2](#), consists of two steps: 1) modeling the across-trial variability of amplitude, latency, and morphology of ERP waveforms ([Mouraux and Iannetti 2008](#); [Spencer 2005](#)) and 2) adding different levels of background EEG noise ([Dien et al. 2005, 2007](#)).

MODELING THE ACROSS-TRIAL VARIABILITY. Similarly to a previous study ([Spencer, 2005](#)), the variability of single-trial ERP amplitude was modeled by multiplying the template with a positive coefficient (Gaussian distribution with 1.5 ± 0.5 , mean $\pm$ SD). The variability of single-trial ERP latency was modeled by shifting the template along the time axis with a latency jitter of ± 20 ms (SD) (Gaussian distribution centered at the peak latency of the group average). The variability of single-trial ERP morphology ([Casarotto et al. 2005](#); [Jung et al. 2001](#)) was modeled by compressing the template along the time axis (Gaussian distribution with 1.5 ± 0.5 , mean $\pm$ SD) centered at the zero-crossing point between the N2 and P2 waves. Each simulated data set consisted of 30 trials. Twelve simulated data sets were generated by varying the amplitude, latency, and morphology of the LEP template by repeating the above three steps ([Fig. 2](#), 1st–3rd panels).

ADDING BACKGROUND NOISE. Real resting EEG data obtained from 12 subjects (30 epochs for each subject, epoch length 1,500 ms, baseline corrected using the prestimulus interval from -500 to 0 ms) were added to the simulated data sets (1 noise trial for each simulated trial) to represent background noise. To assess the performance of the four single-trial methods at different noise levels, the real resting EEG data were multiplied by 11 different weights (ranging from 0.5 to 1.5, with 0.1 as step size) before being added to the simulated data sets. Thus this procedure yielded 11 subsets of data, containing all the different levels of noise for each of the 12 data sets ([Fig. 2](#), 4th and 5th panels).

After a different level of noise was added, the average SNR of the simulated data sets was estimated as follows ([Iyer and Zouridakis 2007](#); [Zouridakis et al. 1997](#)):

$$\text{SNR} = \frac{1}{L} \sum_{i=1}^L \frac{\sigma_{\text{ERP}_i}^2}{\sigma_{\text{EEG}_i}^2} \quad (4)$$

where L is the number of trials, $\sigma_{\text{ERP}_i}^2$ is the variance of the simulated i th ERP signal resulting from the addition of amplitude, latency, and shape variation to the real ERP template (e.g., [Fig. 2](#), 3rd panel), and $\sigma_{\text{EEG}_i}^2$ is the variance of the i th EEG noise (e.g., [Fig. 2](#), 4th panel), which has been merged into the i th ERP signal.

Comparison with other methods of single-trial estimation. The performance of the four methods (MLR, MLR_d, WF+MLR, and WF+MLR_d) was assessed by measuring the correlation coefficient between the original and the estimated single-trial parameters (i.e., the latency, amplitude, and morphology of N2 and P2 of each single trial) ([Jaskowski and Verleger 1999](#)). For each level of noise, the correlation coefficients obtained with the four methods were compared using a two-way repeated-measures analysis of variance (ANOVA) to explore the effects of WF (2 levels: “with WF” and “without WF”) and dispersion (2 levels: “MLR” and “MLR_d”), as well as the possible interaction between these two factors ([Fig. 3](#)). In addition, to test the overall performance of the four methods (i.e., independently of noise level), we also performed the same two-way repeated-measures ANOVA on the correlation coefficients averaged across the 11 noise levels. When significant, post hoc paired t -tests were used to perform pairwise comparisons. All statistical analyses were carried out using SPSS 16.0 (SPSS, Chicago, IL) ([Fig. 3](#)).

To investigate the reasons for possible significant effects of the two experimental factors on the performance of the four applied single-trial approaches, we performed the following two additional analyses. To explore the possible effect of WF, we estimated the SNR of ERPs using [Eq. 4](#). The SNR values of ERPs before and after WF were compared using the nonparametric Wilcoxon test because of their nonnormal distribution.

To explore whether the possible effect of dispersion was simply due to the larger number of regressors used in the MLR_d conditions, we performed an *F*-test, as described by [Motulsky and Christopoulos \(2004\)](#). Obviously, the MLR_d model (which has 2 more basis sets than the MLR model) will always be able to fit the data at least as well as the MLR model (which has fewer parameters). Therefore, any method to compare a simple model with a more complex model has to balance the decrease in sum of squares with the increase in the number of parameters, and this can be achieved by using the *F*-test ([Motulsky and Christopoulos 2004](#)), which can determine whether the MLR_d model gives a significantly better fit to the data than the MLR model.

Single-Trial Estimation of LEP Data

The WF+MLR_d approach, which proved to be the most accurate of the four explored approaches to identify ERP in single trials, was applied to a real LEP data set and compared with the WF+MLR approach.

Subjects. EEG data were collected from 12 healthy volunteers (8 females and 4 males) ages 23–42 yr (29 ± 5, mean ± SD). All participants gave written informed consent, and the local ethics committee approved the procedures.

Nociceptive stimulation. Noxious radiant-heat stimuli were generated by an infrared neodymium-yttrium-aluminum-perovskite laser with a wavelength of 1.34 μm (Electronical Engineering, Florence, Italy). These laser pulses activate directly nociceptive terminals in the most superficial skin layers ([Baumgartner et al. 2005](#); [Iannetti et al. 2006](#)). Laser pulses were directed to the dorsum of the right hand, and a He-Ne laser pointed to the area to be stimulated. The laser pulse was transmitted via an optic fiber and focused by lenses to a spot diameter of ~8 mm (50 mm²) at the target site. The duration of the laser pulses was 4 ms, and its energy was 3.5 J. With these parameters, laser pulses elicit a clear pinprick sensation, related to the activation of Aδ skin nociceptors ([Iannetti et al. 2006](#)). After each stimulus, the laser beam target was shifted by ~20 mm in a random direction, to avoid nociceptor fatigue and sensitization. The laser beam was controlled by a computer that used two servomotors (HS-422; Hitec RCD, Poway, CA; angular speed 60°/160 ms) to orient it along two perpendicular axes ([Lee et al. 2009](#)).

Experimental paradigm and EEG recording. EEG data were collected in a single recording block. Participants were seated in a comfortable chair and wore protective goggles. They were asked to focus their attention on the stimuli, relax their muscles, and keep their eyes open and gaze slightly downward. Acoustic isolation was ensured using earplugs and headphones. Both the laser beam and the controlling motors were completely screened from the view of the participants. The experiment consisted of a single block of 30 trials, with an interstimulus interval ranging between 15 and 18 s. Participants were asked to rate verbally the intensity of the sensation evoked by each stimulus, using a scale ranging from 0 to 10, where 0 was “no pain” and 10 was “pain as bad as it could be” ([Jensen and Karoly 2001](#)).

The EEG was recorded using 30 Ag-AgCl electrodes placed on the scalp according to the International 10-20 system, using the nose as reference. To monitor ocular movements and eye blinks, electrooculographic (EOG) signals were simultaneously recorded from two surface electrodes, one placed over the lower eyelid and the other placed 1 cm lateral to the outer corner of the orbit. The

electrocardiogram was recorded using two electrodes placed on the dorsal aspect of the left and right forearms. Signals were amplified and digitized using a sampling rate of 1,024 Hz and a precision of 12 bits, giving a resolution of 0.195 $\mu\text{V}/\text{digit}$ (System Plus; Micromed, Treviso, Italy).

EEG data preprocessing. EEG data were imported and processed using EEGLAB ([Delorme and Makeig 2004](#)), an open source toolbox running under the MATLAB environment. Continuous EEG data were bandpass filtered between 1 and 30 Hz. EEG epochs were extracted using a window analysis time of 1,500 ms (500 ms prestimulus and 1,000 ms poststimulus) and baseline corrected using the prestimulus time interval. To test the possible bias in the automated single-trial LEP detection method, the same number of trials of resting EEG (3,500 to 5,000 ms poststimulus) were extracted from the data set of each subject.

Trials contaminated by eyeblinks and movements were corrected using an independent component analysis (ICA) algorithm ([Delorme and Makeig 2004](#); [Jung et al. 2001](#); [Makeig et al. 1997](#)). EEG epochs were then visually inspected, and trials contaminated by artifacts due to gross movements were removed. In all data sets, individual eye movements showing a large EOG channel contribution and a frontal scalp distribution were clearly seen in the removed independent components.

After artifact rejection, 357 trials remained for the automated ERP detection, leading to a set of 357 epochs (segmented from -500 to $+1,000$ ms after stimulus onset). A corresponding set of 357 “resting EEG epochs” were obtained by segmenting these same trials from 3,500 to 5,000 ms after stimulus onset and were used for testing a possible detection bias.

Enhancement of LEP SNR. The SNR of LEPs was estimated by dividing the N2 peak amplitude (absolute value) by the standard deviation of the average ERP waveform in the prestimulus interval (-500 to 0 ms) ([Debener et al. 2007](#); [Hu et al. 2010](#); [Spencer 2005](#)). Note that the choice of using two different methods to estimate the SNR of ERPs in simulated and real data was driven by the fact that the actual ERP response is known in the simulated data but is not known in the real data. Thus [Eq. 4](#) will provide an accurate estimation of the SNR in simulated data but not in real data. The SNR values of ERPs before and after WF were compared using the nonparametric Wilcoxon test because of their nonnormal distribution.

Single-trial estimation and correlation analysis. By applying both the WF+MLR and WF+MLR_d approaches to the real LEP data set, we estimated single-trial N2 and P2 latencies and amplitudes, as detailed earlier. In addition, we assessed the distortion of the LEP morphology introduced by across-trial averaging of single-trial waveforms. The “distortion ratio” of N2 and P2 waves was estimated as follows:

$$\text{N2 ratio} = \frac{w_{\text{Na}}}{w'_{\text{Ns}}} \quad (5)$$

$$\text{P2 ratio} = \frac{w_{\text{Pa}}}{w'_{\text{Ps}}} \quad (6)$$

where w_{Na} and w_{Pa} are the widths of the N2 and P2 waves at their half-peak maximum in the average waveform, and w_{Ns} and w_{Ps} are the widths of N2 and P2 waves at their half-peak maximum in the single-trial waveforms ([Fig. 4](#)). Note that the distortion ratio of each ERP wave is proportional to the reciprocal of its width.²

Correlations were performed between 1) the estimated single-trial parameters (N2 and P2 latencies and amplitudes) and the corresponding parameters measured from the average waveforms, 2) the estimated single-trial parameters (N2 and P2 latencies and amplitudes) and the corresponding single-trial N2 and P2 distortion ratios, and 3) the estimated single-trial parameters (N2 and P2 latencies, amplitudes, and distortion ratios) and the corresponding single-trial ratings of perceived pain intensity, as follows. 1) For each subject, single-trial latencies and amplitudes of both N2 and P2 waves were averaged across trials and compared with the corresponding values manually measured from the averaged waveforms. Correlation coefficients and their significance were calculated for each of these values. 2) The correlations between single-trial parameters (N2 and P2 latencies and amplitudes) and the corresponding distortion ratios (i.e., N2 latencies vs. N2 ratios, N2 amplitudes vs. N2 ratios, P2 latencies vs. P2 ratios, P2 amplitudes vs. P2 ratios) were tested using the correlation coefficient for parametric data (Pearson R) for each subject (Iannetti et al. 2005b). The obtained correlation coefficients were transformed to Z values using the Fisher R -to- Z transformation. The Z values were finally compared against zero using a one-sample t -test. 3) The same statistical approach was applied to assess the relationship between single-trial latencies, amplitudes, and distortion ratios of both N2 and P2 waves and the corresponding single-trial ratings of perceived pain intensity.

Detection bias. Finally, to rule out any possible bias in the estimation (i.e., the detection of a signal when there is none), the same WF+MLR_d approach was applied to the resting EEG epochs obtained from the same subjects. This possible detection bias was tested by comparing the obtained N2 and P2 amplitude values against zero using a one-sample t -test (Mayhew et al., 2006; Hu et al., 2010).

RESULTS

Single-Trial Analysis Approaches

Multiple linear regression.

Figure 1, top, shows the N2 and P2 regressors obtained from a representative, WF-filtered LEP waveform and a single trial fitted using MLR. Note that there are two regressors (an average and the temporal derivative) for each ERP wave and that the temporal derivatives allow modeling of the variability of single-trial latency but not the variability of single-trial morphology.

Multiple linear regression with dispersion term.

Figure 1, bottom, shows the N2 and P2 regressors obtained from a representative, WF-filtered LEP waveform and a single trial fitted using the MLR_d based on PCA. Note that the single-trial variability of both latency and morphology is now modeled. The first three PCs explained 98.3 ± 0.3 and $98.6 \pm 0.2\%$ of the variance of the variability matrices for the N2 and P2 waves, respectively. Thus the ERP waves are modeled using three regressors (PC1, PC2, and PC3), and the variations of both latency and morphology are correctly represented by these regressors.

Single-Trial Estimation on Simulated Data Set

Comparison with other methods of single-trial estimation.

Figure 3 shows the performance of the four methods to estimate single-trial ERP parameters (MLR, MLR_d, WF+MLR, and WF+MLR_d), quantitatively assessed by calculating the correlation coefficient between the true and the estimated single-trial values. As expected, when the noise level increased, the correlation coefficients between the true and the estimated values became smaller (Fig. 3, top row). The WF+MLR_d approach yielded the highest correlation coefficients for all the ERP parameters (latency, amplitude, and morphology of N2 and P2). Figure 3, middle row, shows the results of the two-way repeated-measures ANOVA for each noise level. At low SNRs there was a significant main effect of the

factor WF ($P < 0.05$). In contrast, at high SNRs there was a significant main effect of the factor dispersion ($P < 0.05$). In most cases there was no significant interaction between these two factors. When considering the performance of the four methods to estimate single-trial variability of ERP morphology, there was a significant main effect of the factor dispersion at virtually all SNRs.

[Figure 3](#), bottom row, shows the correlation coefficients across all levels of noise. There were significant main effects of the factors WF and dispersion on the latency, amplitude, and morphology of both the N2 and the P2 waves ($P < 0.001$ for both factors and all ERP parameters). There was also a significant interaction between these two factors in modulating the P2 amplitude ($P < 0.001$). Post hoc comparisons revealed that the correlation coefficients obtained using the WF+MLR_d approach were significantly larger than those obtained using either the MLR or the MLR_d approach for all ERP parameters ($P < 0.001$, 2-tailed paired t -test) and also larger than those obtained using the WF+MLR approach on N2 latency, N2 amplitude, N2 morphology, P2 latency, and P2 morphology ($P < 0.001$, 2-tailed paired t -test).

The SNRs of the simulated data sets, which were obtained by adding real resting EEG data with different levels of noise (ranging from 0.5 to 1.5, with 0.1 as step size) to the theoretical data sets, ranged between 0.45 ± 0.22 and 4.0 ± 2.0 . At every level of noise, WF significantly enhanced the SNR ($P < 0.005$, Wilcoxon test).

The better performance of the MLR_d approach, compared with the MLR approach, was not simply due to the larger number of regressors used in the MLR_d conditions but to the fitting of biologically relevant information (i.e., the trial-to-trial variability in LEP morphology). This was demonstrated by the significant result of the F -test performed between the goodness of fit obtained with the two approaches, indicating that MLR_d provides a better fit to the single-trial data [$F(2, 505) = 805$, $P < 0.0001$].

Altogether, the assessment of the performance of the different approaches on the simulated data indicates that WF strongly enhances the SNR of ERPs and provides a more accurate and reliable single-trial estimation, especially when the level of noise is high (low SNR). On the other hand, MLR_d improves the performance of single-trial estimation, especially when the level of noise is low (high SNR), whereas it would fit part of the noise when the SNR of the ERP is low. For these reasons, the WF+MLR_d approach provides the most accurate single-trial estimation compared with the other three approaches, and it was applied to explore the real LEP data set.

Single-Trial Estimation of LEP Data

Single-trial estimation and correlation analysis. The bidimensional plots in [Fig. 5](#), middle, show the effect of WF on single-trial LEP waveforms and resting EEG. Whereas in the LEP waveforms the WF significantly enhanced the SNR of the phase-locked responses (from 7.3 ± 3.3 to 21.2 ± 10.4 , $P = 0.002$, Wilcoxon test), in the resting EEG the WF only reduced the noise. [Figure 5](#), left, shows that only the scalp topographies of the average LEP waveforms at the latency of the N2 and P2 peak showed the expected central distribution maximal at the vertex ([Bromm and Treede 1984](#); [Lee et al. 2009](#); [Mouraux and Iannetti 2008, 2009](#)).

The resulting distortion ratios, estimated using [Eqs. 5](#) and [6](#), were averaged across trials for each subject. The average distortion ratios were significantly >1 (N2 ratio: 1.28 ± 0.15 , $P < 0.001$; P2 ratio: 1.45 ± 0.25 ; $P < 0.001$; 1-sample t -test).

CORRELATION BETWEEN SINGLE TRIALS AND AVERAGE WAVEFORMS.

[Figure 6](#) shows the correlations between the single-trial LEP parameters and the corresponding parameters manually measured from the average waveforms. The averages of single-trial estimates of N2 latency, N2 amplitude, P2 latency, and P2 amplitude values showed a strong correlation with the

corresponding values measured in standard averaged waveforms (N2 latency: $R = 0.9835$, $P < 0.0001$; N2 amplitude: $R = 0.9913$, $P < 0.0001$; P2 latency: $R = 0.9834$, $P < 0.0001$; P2 amplitude: $R = 0.9491$, $P < 0.0001$). The N2 and P2 latency values were almost identical (N2 latency, single trials: 221 ± 28 ms; N2 latency, standard average: 221 ± 32 ms; $P = 0.86$, 2-tailed t -test; P2 latency, single trials: 345 ± 56 ms; P2 latency, standard average: 350 ± 60 ms; $P = 0.15$, 2-tailed t -test). In contrast, the N2 and P2 amplitude values were significantly greater in the single-trial estimates than in the standard averaged waveforms (N2 amplitude, single trials: -20.5 ± 12.2 μ V; N2 amplitude, standard average: -15.4 ± 8.6 μ V; $+33 \pm 10\%$ increase, $P = 0.0017$, 2-tailed t -test; P2 amplitude, single trials: 15.5 ± 7.9 μ V; P2 amplitude, standard average: 10.6 ± 6.1 μ V; $+46 \pm 25\%$ increase, $P < 0.001$, 2-tailed t -test).

CORRELATION BETWEEN SINGLE-TRIAL PARAMETERS.

[Figure 4](#), *bottom* (2nd and 3rd rows), shows the correlations between single-trial parameters (N2 and P2 latencies and amplitudes) and the corresponding single-trial distortion ratios. The latency values and distortion ratios of N2 wave correlated significantly (mean $R = -0.147$, $P = 0.020$), whereas those of P2 wave did not (mean $R = -0.091$, $P = 0.162$). The amplitude values and distortion ratios of both waves displayed strong correlations (N2 wave: mean $R = -0.295$, $P < 0.001$; P2 wave: mean $R = -0.276$, $P < 0.001$).

CORRELATION BETWEEN THE SINGLE TRIAL PARAMETERS AND PERCEIVED PAIN.

[Figure 4](#), *bottom* (1st row), shows the correlations between single-trial distortion ratios of each wave and the corresponding intensity of pain perception. The single-trial N2 distortion ratios showed a significant correlation with the intensity of pain perception (mean $R = -0.191$, $P = 0.029$), whereas the single-trial P2 distortion ratio did not (mean $R = -0.091$, $P = 0.162$). [Figure 7](#) shows the correlations between single-trial LEP parameters, estimated using both WF+MLR and WF+MLR_d approaches, and the corresponding intensity of pain perception. Whereas the single-trial N2 and P2 latencies, estimated using both approaches, did not correlate with the intensity of pain perception (using WF+MLR, N2: mean $R = -0.106$, $P = 0.123$; P2: mean $R = -0.060$, $P = 0.519$; using WF+MLR_d, N2: mean $R = 0.038$, $P = 0.611$; P2: mean $R = -0.107$, $P = 0.067$), both the single-trial N2 and P2 amplitudes significantly correlated with the intensity of pain perception. Notably, the single-trial correlations obtained using WF+MLR_d (N2: mean $R = 0.377$, $P < 0.001$; P2: mean $R = 0.251$, $P = 0.003$) are notably better than those obtained using WF+MLR (N2: mean $R = 0.349$, $P = 0.001$; P2: mean $R = 0.228$, $P = 0.021$).

Detection bias. To test whether the method used to estimate single-trial N2 amplitude and P2 amplitude introduced any detection bias, we applied the same method to an equal number of resting EEG epochs obtained from all subjects. The single-trial estimate of response amplitude from resting EEG epochs yielded a mean ($\pm$ SD) amplitude value of -0.72 ± 10.8 μ V for the N2 fit and 0.36 ± 8.3 μ V for the P2 fit. These amplitude values were not significantly different from zero (N2 amplitude: $P = 0.25$; P2 amplitude: $P = 0.47$, 1-sample t -test). A comparison of the single-trial estimates obtained in the LEP waveforms with the resting EEG epochs in a representative subject is shown in [Fig. 5](#), *right*. This result confirms that the described method provides an unbiased estimate of single-trial N2 and P2 amplitudes.

Comparison with WF+MLR approach.

[Figure 8](#) shows the average of single-trial estimates of N2 latency, N2 amplitude, P2 latency, and P2 amplitude values obtained using the WF+MLR_d and the WF+MLR approaches. As expected, the average of single-trial estimates of N2 and P2 latencies obtained with the two approaches were almost identical (N2 latency estimated with WF+MLR: 223 ± 32 ms; N2 latency estimated with WF+MLR_d: 221 ± 28 ms; $P = 0.36$, 2-tailed t -test; P2 latency estimated with WF+MLR: 342 ± 58 ms; P2 latency estimated with WF+MLR_d: 345 ± 56 ms; $P = 0.38$, 2-tailed t -test). In contrast, the N2 and P2 amplitude values were significantly greater when estimated using the WF+MLR_d approach (N2 amplitude estimated with

WF+MLR: -19.5 ± 11.6 μ V; N2 amplitude estimated with WF+MLR_d: -20.5 ± 12.2 μ V; $+5 \pm 5.8\%$ increase, $P = 0.0029$, 2-tailed t -test; P2 amplitude estimated with WF+MLR: 14 ± 7.9 μ V; P2 amplitude estimated with WF+MLR_d: 15.5 ± 7.9 μ V; $+14 \pm 19\%$ increase, $P < 0.001$, 2-tailed t -test).

DISCUSSION

Using simulated ERP data sets with varying levels of noise, we compared the performance of four approaches to estimate ERPs in single trials. We provide evidence that the WF+MLR_d approach yielded the most accurate estimate of the peak latency and amplitude of the N2 and P2 waves of LEPs. The better performance of the WF+MLR_d approach was due 1) to the WF procedure, which enhances the SNR effectively in both single-trial and average ERP waveforms, and 2) to the inclusion in the model of the information concerning the variability in single-trial ERP morphology. When applied to a real LEP data set, the WF+MLR_d approach provided reliable estimation of single-trial latency and amplitude of the N2 and P2 waves, allowing us to measure the distortion of LEP morphology caused by across-trial averaging and to perform meaningful correlations at the single-trial level, both between the different parameters of LEPs and the intensity of perceived pain and within the different parameters of LEPs.

We observed a strong correlation between the intensity of pain perception and the N2 and P2 single-trial amplitudes; this was best detected when single-trial amplitudes were estimated using WF+MLR_d. In addition, we observed a significant positive correlation between single-trial N2 amplitude and N2 width, between single-trial P2 amplitude and P2 width, and between single-trial N2 width and the intensity of pain perception.

WF to Enhance the SNR of ERPs

Compared with the windowed Fourier transform, the continuous wavelet transform (CWT) offers an optimal compromise for time-frequency resolution and is therefore very suitable for exploring a wide frequency range ([Chorlian et al. 2003](#); [Effern et al. 2000](#); [Mouraux and Iannetti 2008](#); [Mouraux and Plaghki 2004](#); [Quiroga and Garcia 2003](#); [Tognola et al. 1998](#)). In this study, by thresholding the average of single-trial time-frequency representations of single-subject ERPs, we generated a wavelet filter that captures both the phase-locked and non-phase-locked responses and provides a time-varying filter.

In the present study, we have shown that WF successfully enhanced the SNR of both simulated and real ERP data sets ([Fig. 5, middle](#)). Importantly, when the same filter was applied to resting EEG data, this procedure significantly reduced the contribution of noise, without generating spurious ERP responses ([Fig. 5, middle](#)). The lack of detection bias is consequent to the fact that WF preserves, in each single trial, the signal corresponding to the time-frequency representation of the ERP regardless of its phase ([Hu et al. 2010](#)). For this reason, when both the MLR_d and the WF+MLR_d approaches were used, the average amplitude of single-trial estimates of the ERP peaks was negligible (MLR_d: 0.67 ± 13.7 μ V for the N2 fit and -0.97 ± 10.6 μ V for the P2 fit, $P = 0.35$ and $P = 0.16$, respectively; WF+MLR_d: -0.72 ± 10.8 μ V for the N2 fit and 0.36 ± 8.3 μ V for the P2 fit, $P = 0.25$ and $P = 0.47$, respectively), but, crucially, the variances of these estimates were significantly smaller when the WF+MLR_d approach rather than the MLR_d approach was used ($P < 0.001$ for both N2 and P2, test of homogeneity of variances) ([Fig. 5](#)). This important difference is due to the fact that WF removes a large amount of noise and thus reduces the power of resting EEG oscillation.

MLR_d to Detect Single-Trial ERPs

We observed that a multiple linear regression with dispersion term using three basis sets on preliminarily wavelet-filtered data (i.e., the WF+MLR_d approach) provided an improved estimate of the peak latency and amplitude of ERP waves ([Fig. 3](#)). The decision of selecting the first three outputs of the PCA performed on the variability matrices was based on the fact that the first three parameters were

able to explain most of the variability of single-trial ERPs, namely, their latency, amplitude, and morphology (Spencer 2005). To capture the variability of these three parameters, we used a nonparametric approach, based on PCA decomposition, because the real function expression of single-trial EEG responses is unknown. When investigating the similar problem of response recognition in BOLD fMRI data, Hossein-Zadeh et al. (2003) also observed that the first three PCs were able to capture the largest part of response variability. These first three PCs were compared with the three Taylor expansion series (i.e., a basic gamma function, a derivative of the gamma function with respect to time variability and a derivative of the gamma function with respect to shape variability) and were shown to be extremely similar, albeit not identical, to the Taylor expansion series (see Fig. 11 in Hossein-Zadeh et al. 2003). This comparison is important, because in fMRI the real function describing the response (e.g., a gamma function) is known, whereas this is not the case for the ERP response (Boynton et al. 1996). Thus it is likely that also the basis set with three regressors generated by PCA and used in the present study mainly captures the latency jitter and the variability in morphology of ERPs, whereas the variability in amplitude is captured by their coefficients of these basis sets. This results in a better single-trial estimate compared with the estimates obtained without including the variability in morphology in the model (i.e., the MLR approach, Fig. 3).

In addition, we used the estimated single-trial LEP waveforms to show that the across-trial averaging procedure significantly distorts their morphology (N2 and P2 distortion ratios: 1.28 ± 0.15 , $P < 0.001$; 1.45 ± 0.25 , $P < 0.001$, 1-sample *t*-test, respectively). This analysis, performed by calculating a distortion ratio for each single-trial, showed that N2 and P2 waves have a higher frequency in single trials than in averages, as indicated by the fact that the distortion ratios are >1 for both the N2 and the P2 waves. Thus the change in frequency of the N2 and P2 waves caused by across-trial averaging is likely to be due to trial-to-trial latency jitter, which results in a smoother average LEP waveform. Hence, including a dispersion term in the MLR analysis appears necessary, because it allows a more accurate fitting of higher frequency, single-trial waveforms. In addition, the observation of a significant variability in the within-subject distortion ratios (widths) indicates that there is also a trial-to-trial variability in LEP morphology, thus indicating that the inclusion of a dispersion term in the MLR analysis would be desirable even in the absence of trial-to-trial latency jitter, to improve the single-trial fitting and thus obtain a better estimate of latency and amplitude of the N2 and P2 LEP peaks.

In addition, a significant positive correlation between the width of the N2 wave and the intensity of pain perception (as well as a trend of positive correlation between the width of the P2 wave and the intensity of pain perception) (N2 wave: $R = -0.191$, $P = 0.029$; P2 wave: $R = -0.091$, $P = 0.162$) indicated that the width of these ERP waves is physiologically meaningful. Thus the inclusion of a dispersion term in the MLR analysis is important to capture the physiological information reflected in the variability of the morphology of each ERP wave. Accordingly, we observed a better correlation between the intensity of pain perception and the corresponding single-trial N2 and P2 amplitudes estimated using WF+MLR_d than between those estimated using WF+MLR (Fig. 7).

We observed that including a factor capturing the variability in ERP morphology significantly improved the estimate of single-trial parameters when the level of noise in the data was low (Fig. 3). This observation makes sense, because introducing the variability in morphology as a fitting factor when processing noisy data might increase the amount of fitted noise, thus reducing both the sensitivity and the specificity of the MLR_d approach. Indeed, the performance of MLR_d and MLR in extracting single-trial amplitude values was very similar when processing data with a high level of noise (Fig. 3). This implies that the use of MLR_d is beneficial either 1) when the SNR of the raw data is high, per se, or 2) when the SNR of the data is effectively enhanced by preliminarily filtering the data (e.g., a simple bandpass filter or the WF described in this article). Thus the combination of the WF with the MLR_d approach allows to broaden the applicability of an MLR_d-based single-trial estimation, even to responses

whose SNR is typically low (e.g., when analyzing early-latency somatosensory ERPs like the N20 wave or when analyzing simultaneously collected EEG/fMRI data; [Debener et al. 2005, 2006](#); [Iannetti et al. 2005a](#)).

Note that the WF+MLR_d was only applied to one recorded channel (Cz) because 1) WF is computationally demanding and 2) Cz is the scalp channel at which the main LEP waves display a maximal amplitude, and consequently, N2 and P2 are recommended to be measured at Cz in both basic and clinical settings ([Treede et al. 2003](#)).

Crucially, the MLR method allows modeling of the amplitude of each single trial regardless of its phase (i.e., regardless of its positive or negative polarity). Therefore, if a sufficient number of trials is considered, the contribution of non-event-related peaks in the signal will cancel out and tend toward zero. For this reason, when we applied the WF+MLR_d approach to the resting EEG data, although some positive and negative fittings were observed, the average peak amplitudes of the single-trial estimates of both N2 and P2 waves were negligible ($-0.72 \pm 10.8 \mu\text{V}$ for the N2 fit and $0.36 \pm 8.3 \mu\text{V}$ for the P2 fit; N2 amplitude: $P = 0.25$; P2 amplitude: $P = 0.47$, 1-sample t -test) ([Fig. 5, right](#)).

N2 and P2 Peaks in Single-Trial LEPs

When the WF+MLR_d approach was applied to real LEP data, the averages of estimated single-trial N2 and P2 latencies were almost identical to the N2 and P2 latencies measured in the across-trial average waveforms ([Fig. 6, left](#)). In contrast, the averages of estimated single-trial N2 and P2 amplitudes were significantly larger than those measured in the across-trial average waveforms ($+33 \pm 10\%$ and $+46 \pm 25\%$, respectively) ([Fig. 6, right](#)).

The large difference in amplitude between the estimate of ERP amplitude at single-trial level and the measure of ERP amplitude in the average ERP waveform is likely to be due to the across-trial latency jitter, which results in averaging ERP waves with slightly different phases, thus distorting their shape and reducing their peak amplitude. Importantly, the increase in peak amplitude obtained by measuring the response in single trials was significantly smaller ($P = 0.028$, 2-tailed t -test) for the N2 wave than for the P2 wave ($+33 \pm 10\%$ and $+46 \pm 25\%$, respectively). This finding indicates that the latency jitter of the N2 wave may be significantly smaller than the latency jitter of the P2 waves of LEPs. Considering that the N2 wave is more transient than the P2 waves (i.e., the N2 has higher frequency content than the P2), across-trial averaging would distort it more than P2 waves. Hence, if the latency jitter affecting the N2 and P2 waves was similar, the gain in amplitude resulting from measuring the N2 wave in single trials should be greater than the gain in amplitude resulting from measuring the P2 waves in single trials. Therefore, the fact that the opposite was observed indicates that the N2 wave is affected by less latency jitter than the later P2 wave. Accordingly, the across-trial variability (expressed as SD) of the peak latency of the N2 wave (which is a direct measure of response jitter) was 28 ms, whereas that for the P2 wave was 56 ms. Altogether, these findings indicate that the latency jitter of the N2 wave is smaller than that of the P2 wave. This finding is similar to the results obtained by [Michalewski et al. \(1986\)](#) showing that in auditory evoked potentials, the latency variability of the N2 and P3 peaks is larger than the latency variability of the earlier N1 and P2 peaks, consistent with the notion that early-latency, exogenous responses (like the N1 wave of LEPs; [Lee et al. 2009](#)) are less affected by latency jitter than longer latency, endogenous responses (like the N2 and P2 waves of LEPs; [Mouraux and Iannetti 2009](#)).

The notion that the shorter the peak latency, the smaller the latency jitter is also supported by an additional observation that the peak amplitudes of the N2 and P2 waves in single trials were significantly larger when estimated by WF+MLR_d than when they were estimated by WF+MLR ($+5 \pm 5.8\%$ and $+14 \pm 19\%$ for N2 and P2 wave, respectively) ([Fig. 8](#)). For the same line of reasoning reported above, the smaller underestimation of the N2 peak amplitude can be also explained by the smaller jitter of the N2

peak latency.

Correlation With the Intensity of Pain Perception

Similarly to what was observed in previous single-trial LEP studies ([Arendt-Nielsen 1994](#); [Iannetti et al. 2005b](#)), the peak amplitude of both the N2 and the P2 waves showed a strong positive correlation with the intensity of pain perception ([Fig. 7, top](#)). The strong correlations observed between N2 and P2 peak amplitudes and the intensity of pain perception may have a peripheral explanation, i.e., it can be due to the fact that stronger inputs activate more cortical neurons synchronously, resulting in ERPs of larger amplitudes ([Iannetti et al. 2005b](#)). Indeed, despite the fact that the energy of the laser stimulation is kept constant, the intensity of the afferent input along peripheral and central nociceptive pathways is expected to vary significantly from trial to trial. This variability can be explained not only by changes in the number of nociceptors activated by each delivered laser pulse (because the irradiated skin spot must be changed after each stimulus) ([Arendt-Nielsen and Chen 2003](#)) but also by within-subject variations in receptor density and skin pigmentation ([Plaghki and Mouraux 2003](#)). This variability of the peripheral input can result in trial-to-trial variations in the intensity of pain perception and thus be reflected by the trial-to-trial variations in peak amplitude of the N2 and P2 LEP waves.

In addition, the correlations between N2 and P2 peak amplitudes and the intensity of pain perception are likely to be also contributed by central factors, independent of the variability of the incoming sensory input. Indeed, we recently showed that the N2 and P2 waves represent a late stage of cortical processing related to the perceptual outcome of the nociceptive input ([Lee et al. 2009](#)). Thus a wide range of cognitive factors, such as attention paid to the stimulus and vigilance, which are known to influence both pain perception and the magnitude of the N2 and P2 waves of LEPs ([Legrain et al. 2002](#); [Lorenz and Garcia-Larrea 2003](#)), could have contributed to the observed correlation between N2 and P2 magnitude and the intensity of perception.

Conclusions and Future Directions

Using both simulated and real ERP data sets, we have shown that WF significantly enhances the SNR of ERPs in single trials and that MLR_d effectively captures and measures the variability in the morphology of single-trial ERPs. Therefore, their combination (WF+MLR_d) provides a reliable and unbiased estimate of single-trial ERP latency, amplitude, and morphology and thus has the potential of yielding novel physiological information (e.g., estimating the latency jitter and the correlation with pain perception of different LEP waves). ICA, which is a blind source separation technique that works as a spatial filter and is able to exploit the spatial information contained in multichannel EEG recordings ([Makeig et al. 2004](#)), has been demonstrated to be effective in isolating stimulus-related and ongoing components in single-trial EEG signals ([Debener et al. 2006](#)). Thus WF+MLR_d could also be applied on ICA-filtered EEG signal or directly on stimulus-related ICs to provide better single-trial estimation on higher SNR and spatial-distinct neural activities.

WF+MLR_d also can be applied to estimate single-trial parameters of short-latency evoked potentials (e.g., somatosensory, visual, and auditory evoked potentials) and stimulus-induced EEG responses in the time-frequency domain. Importantly, WF+MLR_d can be applied to explore the dynamics between different features of ERPs, behavioral variables, and measures of brain activity sampled using different neuroimaging techniques and hence may provide new insights into the functional significance of the brain processes underlying the brain responses to sensory stimuli. Finally, this WF+MLR_d approach is also useful in clinical practice. For example, LEPs, which often display longer latency, reduced amplitude, and/or increased wave width in the presence of lesions of the peripheral or central nervous system, are highly recommended to evaluate the function of nociceptive pathways in patients ([Treede et al. 2003](#)). However, signals recorded from patients usually have lower SNR than those recorded from

healthy subjects, thus suggesting the value of using WF+MLR_d to increase the SNR of ERP waveforms, not only in average but also in single trials. Furthermore, the ability to quantify the variability of single-trial amplitude, latency, and morphology will allow a more complete use of the information contained in ERPs in clinical studies.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

L.H., A.M., and G.D.I. conception and design of research; L.H. and M.L. performed experiments; L.H. analyzed data; L.H., M.L., A.M., R.G.W., and G.D.I. interpreted results of experiments; L.H. and G.D.I. prepared figures; L.H. and G.D.I. drafted manuscript; L.H., M.L., A.M., R.G.W., Y.H., and G.D.I. edited and revised manuscript; L.H., M.L., A.M., R.G.W., Y.H., and G.D.I. approved final version of manuscript.

ACKNOWLEDGMENTS

L. Hu is supported by the Lee Wing Tat Medical Research Fund. R. G. Wise is supported by the Medical Research Council. G. D. Iannetti is a University Research Fellow of The Royal Society and acknowledges the support of the Biotechnology and Biological Sciences Research Council.

Footnotes

¹Additional evidence in support of the functional significance of each of the three PCs is provided in the Appendix available online at the IannettiLab website (<http://iannettilab.webnode.com>).

²Additional evidence in support of the accuracy of single-trial morphology estimation is provided in the Appendix available online at the IannettiLab website (<http://iannettilab.webnode.com>).

REFERENCES

- Arendt-Nielsen L. Characteristics, detection, and modulation of laser-evoked vertex potentials. *Acta Anaesthesiol Scand Suppl* 101: 7–44, 1994. [PubMed: 8023675]
- Arendt-Nielsen L, Chen AC. Lasers and other thermal stimulators for activation of skin nociceptors in humans. *Neurophysiol Clin* 33: 259–268, 2003. [PubMed: 14678840]
- Baumgartner U, Cruccu G, Iannetti GD, Treede RD. Laser guns and hot plates. *Pain* 116: 1–3, 2005. [PubMed: 15936889]
- Boynton GM, Engel SA, Glover GH, Heeger DJ. Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16: 4207–4221, 1996. [PubMed: 8753882]
- Bromm B, Treede RD. Laser-evoked cerebral potentials in the assessment of cutaneous pain sensitivity in normal subjects and patients. *Rev Neurol (Paris)* 147: 625–643, 1991. [PubMed: 1763252]
- Bromm B, Treede RD. Nerve fibre discharges, cerebral potentials and sensations induced by CO₂ laser stimulation. *Hum Neurobiol* 3: 33–40, 1984. [PubMed: 6330009]
- Carmon A, Mor J, Goldberg J. Evoked cerebral responses to noxious thermal stimuli in humans. *Exp Brain Res* 25: 103–107, 1976. [PubMed: 1269555]
- Casarotto S, Bianchi AM, Cerutti S, Chiarenza GA. Dynamic time warping in the analysis of event-related potentials. *IEEE Eng Med Biol Mag* 24: 68–77, 2005. [PubMed: 15709539]
- Chorlian DB, Porjesz B, Begleiter H. Determination of human EEG alpha entrainment ERD/ERS using the continuous complex wavelet transform. *Proc Soc Photo Opt Instrum Eng* 5102: 145–155, 2003.
- Cruccu G, Aminoff MJ, Curio G, Guerit JM, Kakigi R, Mauguiere F, Rossini PM, Treede RD, Garcia-Larrea L. Recommendations for the clinical use of somatosensory-evoked potentials. *Clin*

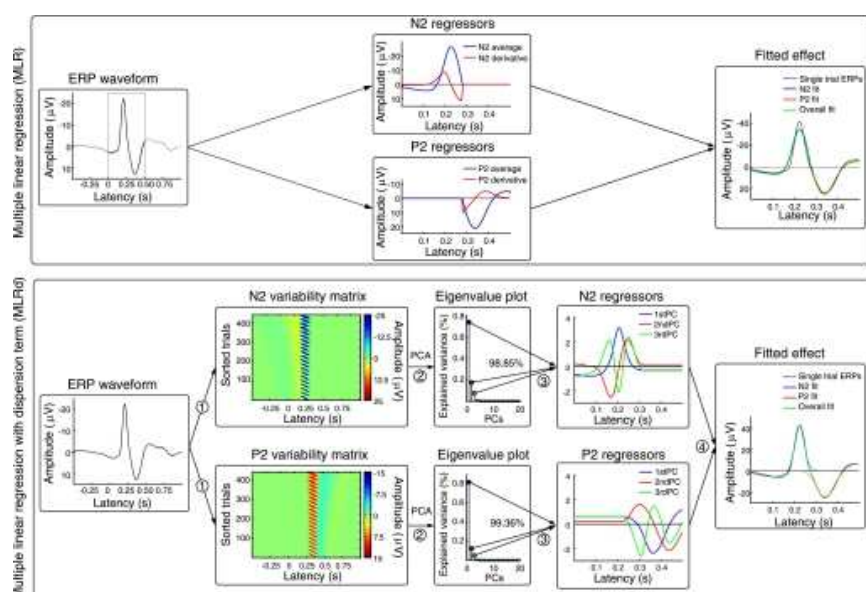
- Neurophysiol 119: 1705–1719, 2008. [PubMed: 18486546]
- Dawson GD. A summation technique for detecting small signals in a large irregular background. J Physiol 115: 2p–3p, 1951.
- Dawson GD. A summation technique for the detection of small evoked potentials. Electroencephalogr Clin Neurophysiol 6: 65–84, 1954. [PubMed: 13141922]
- Debener S, Strobel A, Sorger B, Peters J, Kranczioch C, Engel AK, Goebel R. Improved quality of auditory event-related potentials recorded simultaneously with 3-T fMRI: removal of the ballistocardiogram artefact. Neuroimage 34: 587–597, 2007. [PubMed: 17112746]
- Debener S, Ullsperger M, Siegel M, Engel AK. Single-trial EEG-fMRI reveals the dynamics of cognitive function. Trends Cogn Sci 10: 558–563, 2006. [PubMed: 17074530]
- Debener S, Ullsperger M, Siegel M, Fiehler K, von Cramon DY, Engel AK. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. J Neurosci 25: 11730–11737, 2005. [PubMed: 16354931]
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134: 9–21, 2004. [PubMed: 15102499]
- Dien J, Beal DJ, Berg P. Optimizing principal components analysis of event-related potentials: matrix type, factor loading weighting, extraction, and rotations. Clin Neurophysiol 116: 1808–1825, 2005. [PubMed: 15996897]
- Dien J, Khoe W, Mangun GR. Evaluation of PCA and ICA of simulated ERPs: Promax vs. Infomax rotations. Hum Brain Mapp 28: 742–763, 2007. [PubMed: 17133395]
- Effern A, Lehnertz K, Grunwald T, Fernandez G, David P, Elger CE. Time adaptive denoising of single trial event-related potentials in the wavelet domain. Psychophysiology 37: 859–865, 2000. [PubMed: 11117466]
- Friman O, Borga M, Lundberg P, Knutsson H. Adaptive analysis of fMRI data. Neuroimage 19: 837–845, 2003. [PubMed: 12880812]
- Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R. Event-related fMRI: characterizing differential responses. Neuroimage 7: 30–40, 1998. [PubMed: 9500830]
- Garcia-Larrea L, Frot M, Valeriani M. Brain generators of laser-evoked potentials: from dipoles to functional significance. Neurophysiol Clin 33: 279–292, 2003. [PubMed: 14678842]
- Hossein-Zadeh GA, Ardekani BA, Soltanian-Zadeh H. A signal subspace approach for modeling the hemodynamic response function in fMRI. Magn Reson Imaging 21: 835–843, 2003. [PubMed: 14599533]
- Hu L, Mouraux A, Hu Y, Iannetti GD. A novel approach for enhancing the signal-to-noise ratio and detecting automatically event-related potentials (ERPs) in single trials. Neuroimage 50: 99–111, 2010. [PubMed: 20004255]
- Iannetti GD, Niazy RK, Wise RG, Jezzard P, Brooks JC, Zambreanu L, Vennart W, Matthews PM, Tracey I. Simultaneous recording of laser-evoked brain potentials and continuous, high-field functional magnetic resonance imaging in humans. Neuroimage 28: 708–719, 2005a. [PubMed: 16112589]
- Iannetti GD, Truini A, Galeotti F, Romaniello A, Manfredi M, Cruccu G. Usefulness of dorsal laser evoked potentials in patients with spinal cord damage: report of two cases. J Neurol Neurosurg Psychiatry 71: 792–794, 2001. [PMCID: PMC1737660] [PubMed: 11723205]
- Iannetti GD, Truini A, Romaniello A, Galeotti F, Rizzo C, Manfredi M, Cruccu G. Evidence of a specific spinal pathway for the sense of warmth in humans. J Neurophysiol 89: 562–570, 2003. [PubMed: 12522202]
- Iannetti GD, Zambreanu L, Cruccu G, Tracey I. Operculoinular cortex encodes pain intensity at the earliest stages of cortical processing as indicated by amplitude of laser-evoked potentials in

- humans. *Neuroscience* 131: 199–208, 2005b. [PubMed: 15680703]
- Iannetti GD, Zambreanu L, Tracey I. Similar nociceptive afferents mediate psychophysical and electrophysiological responses to heat stimulation of glabrous and hairy skin in humans. *J Physiol* 577: 235–248, 2006. [PMCID: PMC2000695] [PubMed: 16973704]
- Iyer D, Zouridakis G. Single-trial evoked potential estimation: comparison between independent component analysis and wavelet denoising. *Clin Neurophysiol* 118: 495–504, 2007. [PubMed: 17188929]
- Jaskowski P, Verleger R. Amplitudes and latencies of single-trial ERP's estimated by a maximum-likelihood method. *IEEE Trans Biomed Eng* 46: 987–993, 1999. [PubMed: 10431464]
- Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: *Handbook of Pain Assessment*, edited by Turk DC, Melzack R, editors. New York: Guilford, 2001, p. 15–34.
- Jolliffe IT. *Principal Component Analysis*. New York: Springer, 2002.
- Jung TP, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ. Analysis and visualization of single-trial event-related potentials. *Hum Brain Mapp* 14: 166–185, 2001. [PubMed: 11559961]
- Lee MC, Mouraux A, Iannetti GD. Characterizing the cortical activity through which pain emerges from nociception. *J Neurosci* 29: 7909–7916, 2009. [PubMed: 19535602]
- Legrain V, Bruyer R, Guerit JM, Plaghki L. Nociceptive processing in the human brain of infrequent task-relevant and task-irrelevant noxious stimuli. A study with event-related potentials evoked by CO₂ laser radiant heat stimuli. *Pain* 103: 237–248, 2003. [PubMed: 12791430]
- Legrain V, Guerit JM, Bruyer R, Plaghki L. Attentional modulation of the nociceptive processing into the human brain: selective spatial attention, probability of stimulus occurrence, and target detection effects on laser evoked potentials. *Pain* 99: 21–39, 2002. [PubMed: 12237181]
- Lorenz J, Garcia-Larrea L. Contribution of attentional and cognitive factors to laser evoked brain potentials. *Neurophysiol Clin* 33: 293–301, 2003. [PubMed: 14678843]
- Makeig S, Debener S, Onton J, Delorme A. Mining event-related brain dynamics. *Trends Cogn Sci* 8: 204–210, 2004. [PubMed: 15120678]
- Makeig S, Jung TP, Bell AJ, Ghahremani D, Sejnowski TJ. Blind separation of auditory event-related brain responses into independent components. *Proc Natl Acad Sci USA* 94: 10979–10984, 1997. [PMCID: PMC23551] [PubMed: 9380745]
- Makeig S, Westerfield M, Jung TP, Enghoff S, Townsend J, Courchesne E, Sejnowski TJ. Dynamic brain sources of visual evoked responses. *Science* 295: 690–694, 2002. [PubMed: 11809976]
- Mayhew SD, Dirckx SG, Niazy RK, Iannetti GD, Wise RG. EEG signatures of auditory activity correlate with simultaneously recorded fMRI responses in humans. *Neuroimage* 49: 849–864, 2010. [PubMed: 19591945]
- Mayhew SD, Iannetti GD, Woolrich MW, Wise RG. Automated single-trial measurement of amplitude and latency of laser-evoked potentials (LEPs) using multiple linear regression. *Clin Neurophysiol* 117: 1331–1344, 2006. [PubMed: 16644270]
- Michalewski HJ, Prasher DK, Starr A. Latency variability and temporal interrelationships of the auditory event-related potentials (N1, P2, N2, and P3) in normal subjects. *Electroencephalogr Clin Neurophysiol* 65: 59–71, 1986. [PubMed: 2416547]
- Motulsky H, Christopoulos A. *Fitting Models to Biological Data Using Linear and Nonlinear Regression: a Practical Guide to Curve Fitting*. New York: Oxford University Press, 2004.
- Mouraux A, Iannetti GD. Across-trial averaging of event-related EEG responses and beyond. *Magn Reson Imaging* 26: 1041–1054, 2008. [PubMed: 18479877]
- Mouraux A, Iannetti GD. Nociceptive laser-evoked brain potentials do not reflect nociceptive-specific neural activity. *J Neurophysiol* 101: 3258–3269, 2009. [PubMed: 19339457]
- Mouraux A, Plaghki L. Single-trial detection of human brain responses evoked by laser activation of

- Adelta-nociceptors using the wavelet transform of EEG epochs. *Neurosci Lett* 361: 241–244, 2004. [PubMed: 15135938]
- Nunez PL, Srinivasan R. *Electric Fields of the Brain: the Neurophysics of EEG*. Oxford: Oxford University Press, 2006.
- Orssaud C. Leber's hereditary optic neuropathy. *Orphanet Encyclopedia*, November 2003.
- Pelosi L, Geesken JM, Holly M, Hayward M, Blumhardt LD. Working memory impairment in early multiple sclerosis. Evidence from an event-related potential study of patients with clinically isolated myelopathy. *Brain* 120: 2039–2058, 1997. [PubMed: 9397020]
- Plaghki L, Mouraux A. How do we selectively activate skin nociceptors with a high power infrared laser? Physiology and biophysics of laser stimulation. *Neurophysiol Clin* 33: 269–277, 2003. [PubMed: 14678841]
- Purves AM, Boyd SG. Time-shifted averaging for laser evoked potentials. *Electroencephalogr Clin Neurophysiol* 88: 118–122, 1993. [PubMed: 7681752]
- Quiroga RQ, Garcia H. Single-trial event-related potentials with wavelet denoising. *Clin Neurophysiol* 114: 376–390, 2003. [PubMed: 12559247]
- Sauseng P, Klimesch W, Gruber WR, Hanslmayr S, Freunberger R, Doppelmayr M. Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. *Neuroscience* 146: 1435–1444, 2007. [PubMed: 17459593]
- Spencer KM. Averaging, detection, and classification of single-trial ERPs. In: *Event-Related Potentials: a Methods Handbook*, edited by Handy TC, editor. Cambridge, MA: MIT Press, 2005, p. 209–227.
- Tognola G, Grandori F, Ravazzani P. Wavelet analysis of click-evoked otoacoustic emissions. *IEEE Trans Biomed Eng* 45: 686–697, 1998. [PubMed: 9609934]
- Treede RD, Lorenz J, Baumgartner U. Clinical usefulness of laser-evoked potentials. *Neurophysiol Clin* 33: 303–314, 2003. [PubMed: 14678844]
- Treede RD, Meyer RA, Campbell JN. Myelinated mechanically insensitive afferents from monkey hairy skin: heat-response properties. *J Neurophysiol* 80: 1082–1093, 1998. [PubMed: 9744923]
- Woolrich MW, Behrens TE, Smith SM. Constrained linear basis sets for HRF modelling using Variational Bayes. *Neuroimage* 21: 1748–1761, 2004. [PubMed: 15050595]
- Zouridakis G, Jansen BH, Boutros NN. A fuzzy clustering approach to EP estimation. *IEEE Trans Biomed Eng* 44: 673–680, 1997. [PubMed: 9254981]

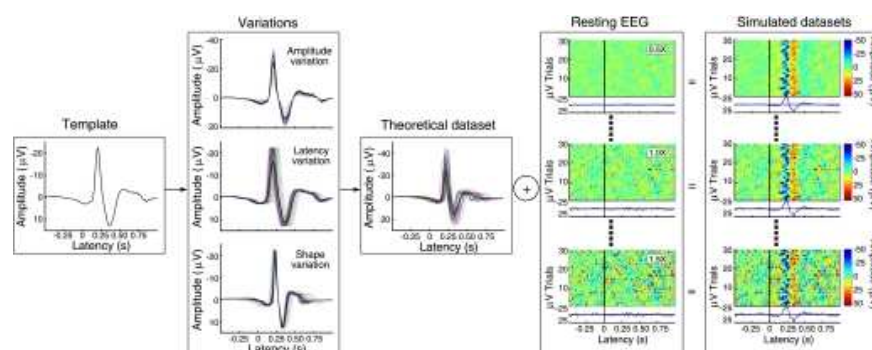
Figures and Tables

Fig. 1.



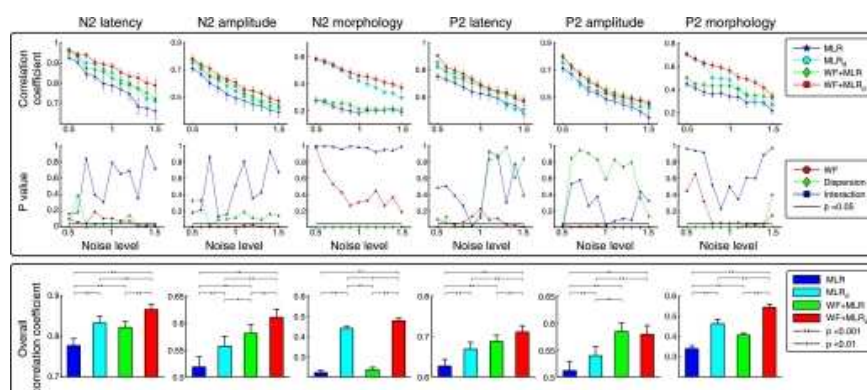
Flowcharts describing the procedures of multiple linear regression (MLR) and multiple linear regression with dispersion term (MLR_d) to estimate the single-trial latency and amplitude of event-related potential (ERP) peaks. *Top*: in MLR, 2 regressors (an average and its temporal derivative) for each ERP peak are calculated from the average ERP waveform within a given poststimulus interval (in this case, 0 to 0.5 s) for each subject. These regressors (in this case, 4) are then applied against each single trial within the same poststimulus interval and used to model each single-trial ERP peak. *Bottom*: in MLR_d, variability matrices that capture the variations of latency and morphology of each ERP peak are generated by simultaneously shifting and compressing the average ERP waveform (*step 1*). These variability matrices, whose order of trials (with the latency shifted and the morphology varied simultaneously) is of no importance, are fed to a principal component analysis (PCA; *step 2*). The resulting 3 main principal components (PCs) are used to define 3 regressors for each peak within a given poststimulus interval (in this case, 0 to 0.5 s; *step 3*). These regressors (in this case, 6) are then applied against each single trial within the same poststimulus interval and used to model each single-trial ERP peak (*step 4*). Note the difference in the fit of the same representative single trial obtained using MLR and MLR_d (*top* and *bottom, far right*), because adding a third regressor (MLR_d) allows modeling the variability in ERP morphology, which is not taken into account in nondispersed MLR.

Fig. 2.



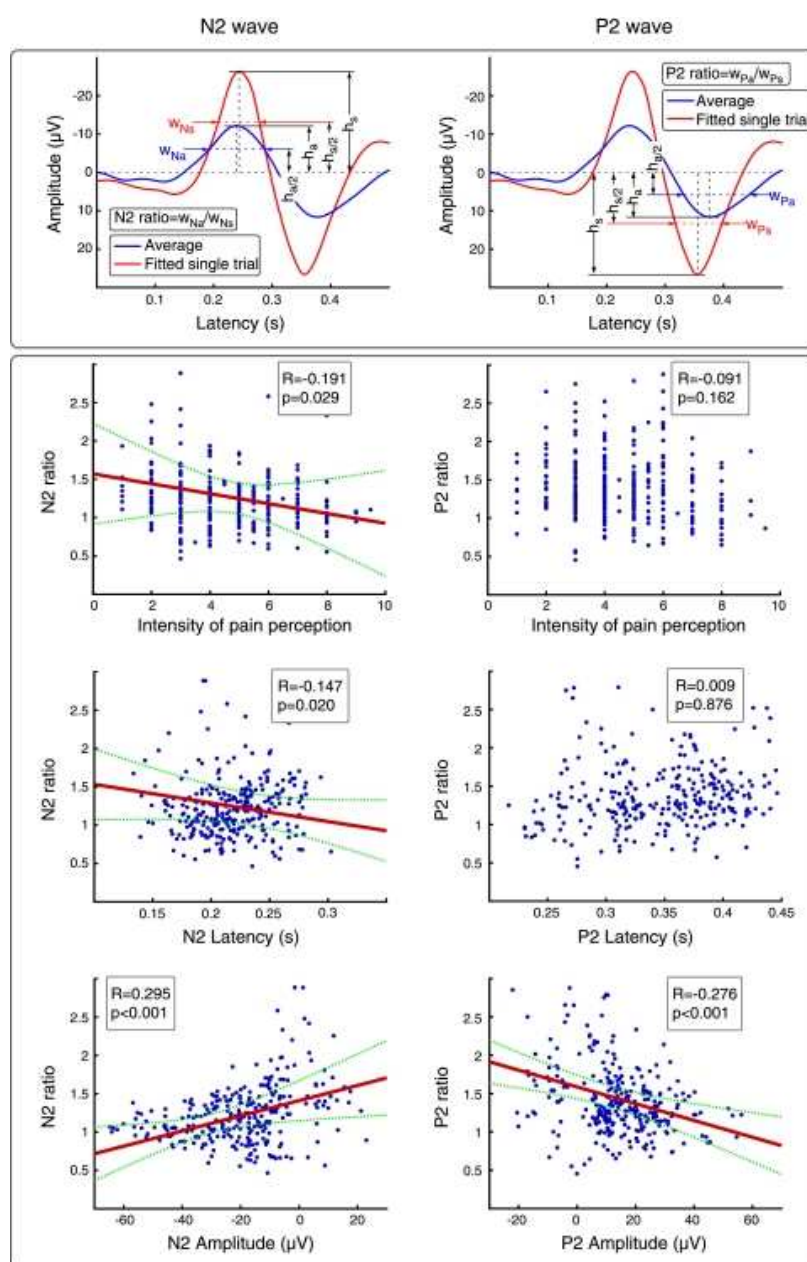
Flowchart describing generation of the simulated data sets with different levels of noise. The ERP group average waveform recorded at channel Cz was used as template (1st panel). Variations in amplitude, latency, and shape were randomly embodied in the ERP template, generating, for each subject, 30 ERP trials (represented by colored waveforms) for each kind of variation (2nd panel). The integration of all these variations into the ERP template generated the simulated ERP data set composed by 30 trials (3rd panel). Eleven levels of noise were generated by multiplying a real resting EEG epoch from that subject by a range of factors (from 0.5 to 1.5, with 0.1 as step size) (4th panel). Finally, each of these 11 noise epochs was added to each of the 30 simulated data sets, thus generating 330 ERP epochs for each subject (5th panel). Consequently, each of the resulting simulated ERP data sets had a different level of signal-to-noise ratio (SNR).

Fig. 3.



Performance evaluation of the 4 methods [MLR, MLR_d, wavelet filtering (WF) + MLR, and WF+MLR_d] to estimate single-trial ERP parameters. *Top row:* correlation coefficients between single-trial estimates and the corresponding single-trial true values at each of the 11 levels of noise. *Middle row:* the effect of WF and dispersion factors was tested using 2-way repeated-measures ANOVA. Note that there was a significant main effect of the factor WF at low SNRs and a significant main effect of the factor dispersion at high SNRs, as well as a significant main effect of the factor dispersion at virtually all SNRs when estimating the single-trial variability of ERP morphology. *Bottom row:* statistical comparison of the overall correlation coefficients (i.e., across all level of noise) obtained with the 4 methods. The effect of WF and Dispersion factors was tested using 2-way repeated-measures ANOVA. Note that the overall correlation coefficients obtained using the WF+MLR_d approach were generally significantly larger than those obtained using the other 3 approaches.

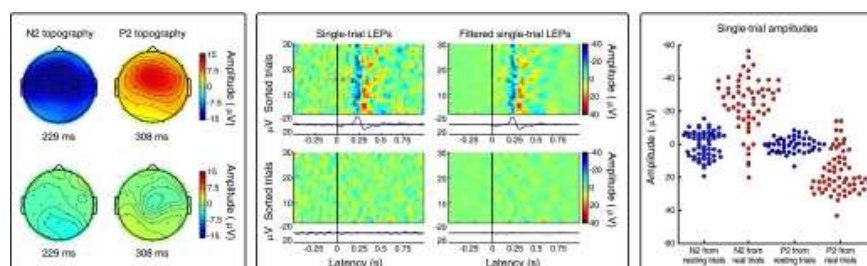
Fig. 4.



Estimation of distortion between single-trial and average ERP waveforms and correlations between single-trial distortion ratio and other single-trial variables. *Top*: the blue waveform represents the average ERP, and the red waveform represents a single-trial ERP. The distortion was estimated by calculating the ratio between the widths at the half-peak height ($h_a/2$ and $h_s/2$) of the average waveform (w_{Na} and w_{Pa}) and of each single trial (w_{Ns} and w_{Ps}). The average distortion ratios of the N2 wave (*left*) and the P2 wave (*right*) were both significantly >1 (N2 ratio: 1.28 ± 0.15 , $P < 0.001$; P2 ratio: 1.45 ± 0.25 , $P < 0.001$, 1-sample t -test). *Bottom*: the single-trial N2 distortion ratios showed a significant negative correlation with the corresponding intensity of pain perception (mean $R = -0.191$, $P = 0.029$), whereas the single-trial P2 distortion ratios did not (mean $R = -0.091$, $P = 0.162$) (1st row). Single-trial N2 distortion ratios showed a significant negative correlation with the corresponding N2 latencies (mean $R = -0.147$, $P = 0.020$), whereas single-trial P2 distortion ratios did not significantly correlate

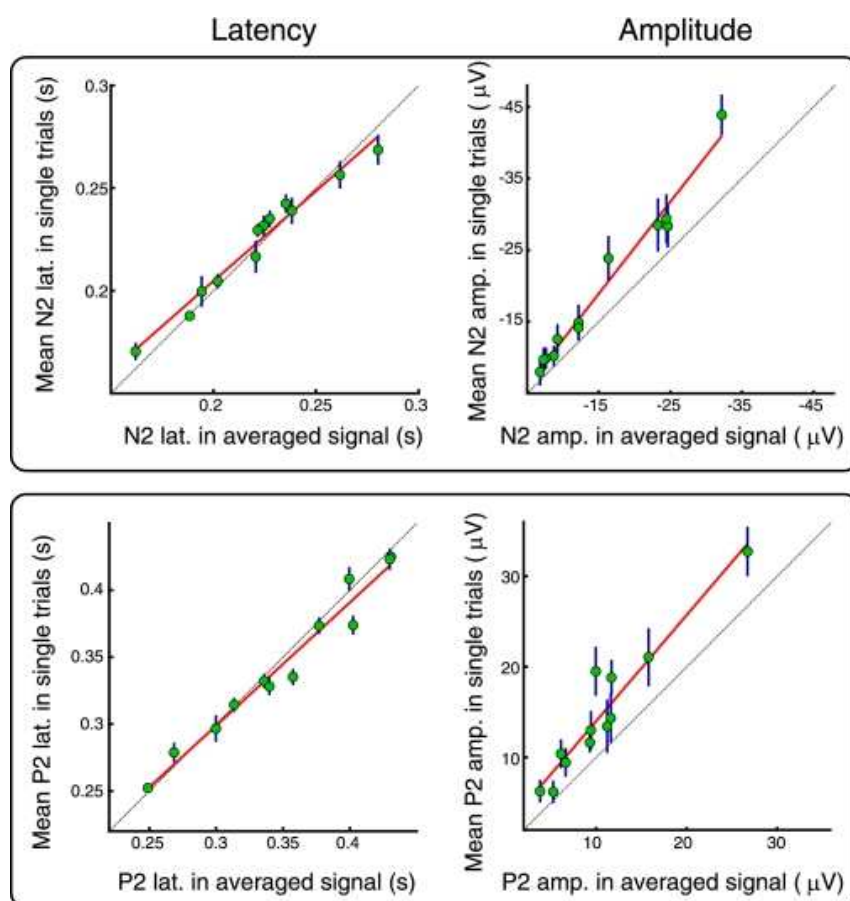
with the corresponding P2 latencies (mean $R = 0.009$, $P = 0.876$). Both the single-trial N2 and P2 distortion ratios significantly correlated with their corresponding single-trial amplitudes (N2 wave: mean $R = -0.295$, $P < 0.001$; P2 wave: mean $R = -0.276$, $P < 0.001$).

Fig. 5.



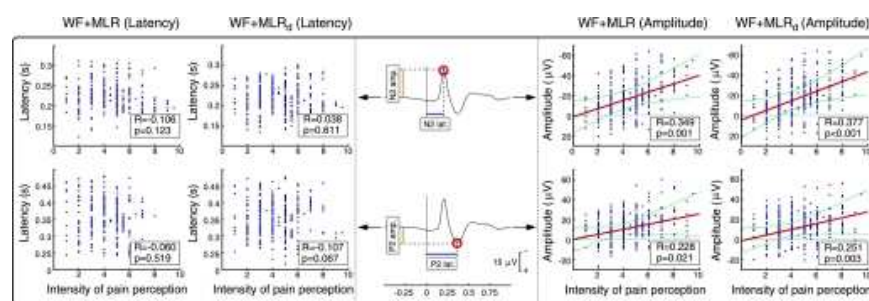
Unbiased automatic detection of single-trial laser-evoked potentials (LEPs) using WF+MLR_d. *Left:* scalp topographies at N2 and P2 peak latencies of the average waveform obtained from LEP trials (*top*) and resting EEG trials (*bottom*). Whereas the topographies of LEP trials were centrally distributed and maximal at the vertex, the topographies of resting EEG trials were randomly distributed. *Middle:* comparison of the effect of WF on LEP trials (*top*) and resting EEG trials (*bottom*). Note that the phase-locked responses in LEP trials were preserved, whereas the noise in both LEP trials and resting EEG trials was markedly reduced. The single trials are sorted according to the order of stimulus presentation. *Right:* single-trial N2 and P2 amplitudes (same data shown at *left* and *middle*) estimated in LEP trials (red) and resting EEG trials (blue). The average N2 and P2 peak amplitudes were -20.5 ± 12.2 and 15.5 ± 7.9 μV in LEP trials and -0.72 ± 10.8 and 0.36 ± 8.3 μV in resting EEG trials. Note that the amplitude values obtained from the resting EEG trials were not significantly different from zero (N2 amplitude: $P = 0.25$; P2 amplitude: $P = 0.47$, 1-sample t -test).

Fig. 6.



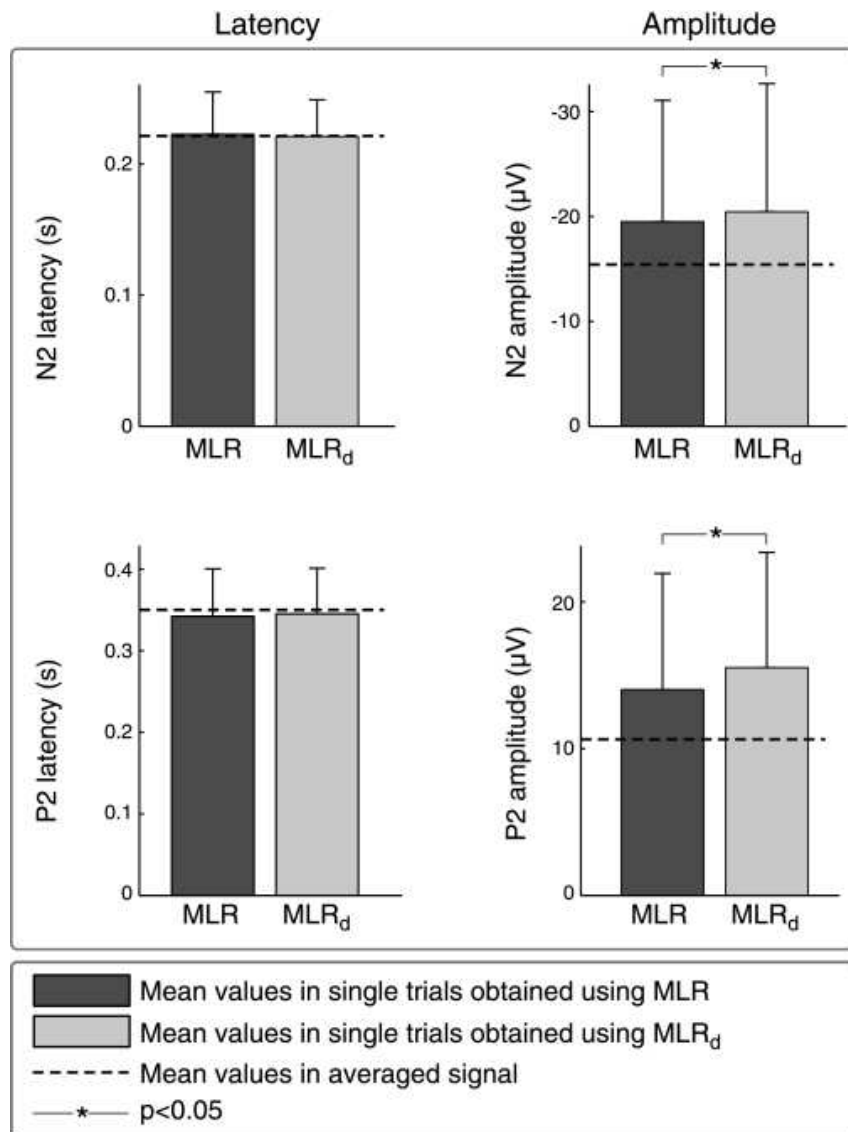
Correlations between the average N2 and P2 latencies and amplitudes estimated using the WF+MLR_d approach and the average N2 and P2 latencies and amplitudes obtained manually from the averaged LEP waveforms. Significant correlations were observed when examining all LEP parameters (N2 latency: $R = 0.9835$, $P < 0.0001$; N2 amplitude: $R = 0.9913$, $P < 0.0001$; P2 latency: $R = 0.9834$, $P < 0.0001$; P2 amplitude: $R = 0.9491$, $P < 0.0001$). Each point represents the data from 1 subject. Vertical error bars represent, for each subject, the variability across trials (expressed as SE). Black dashed lines represent the identity lines, and red solid lines represent the best linear fit. Note that the N2 latency and P2 latency values obtained from the 2 measurements are almost identical (N2 latency: $P = 0.86$; P2 latency: $P = 0.15$, 2-tailed t -test). In contrast, the N2 amplitude and P2 amplitude values were significantly greater in the single-trial estimates than in the standard averaged waveforms (N2 amplitude: +33% increase, $P = 0.0017$; P2 amplitude: +46% increase, $P < 0.001$, 2-tailed t -test). lat., Latency; amp., amplitude.

Fig. 7.



Correlations between single-trial LEP parameters, estimated using both WF+MLR and WF+MLR_d, and intensity of pain perception. *Left:* correlations between single-trial LEP peak latencies and pain perception. Single-trial N2 latencies and P2 latencies did not correlate with the intensity of pain perception, estimated using either WF+MLR (N2: mean $R = -0.106$, $P = 0.123$; P2: mean $R = -0.060$, $P = 0.519$) or WF+MLR_d (N2: mean $R = 0.038$, $P = 0.611$; P2: mean $R = -0.107$, $P = 0.067$). *Right:* correlations between single-trial LEP peak amplitudes and pain perception. When estimated using WF+MLR, both the single-trial N2 amplitudes and P2 amplitudes significantly correlated with the corresponding intensity of pain perception (N2: mean $R = 0.349$, $P = 0.001$; P2: mean $R = 0.228$, $P = 0.021$). When estimated using WF+MLR_d, single-trial N2 amplitudes and P2 amplitudes correlated even more strongly with the corresponding intensity of pain perception (N2: mean $R = 0.377$, $P < 0.001$; P2: mean $R = 0.251$, $P = 0.003$).

Fig. 8.



Comparison of single-trial estimates obtained using the WF+MLR (dark gray) and the WF+MLR_d approaches (light gray). The average of single-trial estimates of N2 and P2 latencies obtained with the 2 approaches were almost identical (N2: $P = 0.36$; P2: $P = 0.38$, 2-tailed t -test). In contrast, the average of single-trial estimates of N2 and P2 amplitudes were significantly larger when the WF+MLR_d approach was used (N2: +5%, $P = 0.0029$; P2: +14%, $P < 0.001$, 2-tailed t -test). Vertical error bars represent the variability across subjects (expressed as SD). Dashed lines represent the mean values of N2 and P2 latencies or amplitudes measured manually from the across-trial average waveforms.