Psychophysiology, *41* (2004), 845–860. Blackwell Publishing Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2004.00243.x

The application of jackknife-based onset detection of lateralized readiness potential in correlative approaches

JUTTA STAHL AND HENNING GIBBONS
Georg Elias Mueller Institute for Psychology, University of Goettingen, Goettingen, Germany

Abstract

The onset of the lateralized readiness potential (LRP) represents an interesting parameter within mental chronometry. According to Miller, Patterson, and Ulrich (1998), reliability increases when LRP onsets are detected in jackknifed LRP waveforms. Until now, jackknifed LRP onsets could be analyzed in factorial designs only. The aim of the present study was to extend the application of jackknifing to correlative approaches (e.g., in personality research). Using different onset scoring techniques, in several simulations run on realistic LRP data jackknifing and single-subject procedures were compared regarding their estimates of a simulated true correlation. For all scoring techniques but regression, jackknifed coefficients were on average higher than the respective single-subject coefficients, and statistical power was higher. Overall, combining jackknifing with a 50%-relative onset criterion is recommended.

Descriptors: Onset of lateralized readiness potential, Jackknifing procedure, Simulation study, Correlation coefficient, Relative criterion

In recent psychophysiological studies using the methodology of the electroencephalogram (EEG), the so-called lateralized readiness potential (LRP) has become a helpful tool to explore information processing and cognitive functions in choice reaction time (RT) tasks (Coles, 1989). Research on mental chronometry (Posner, 1978) has benefited from the use of the LRP as an additional measure besides RT to investigate the information transmission between, and time course of, different cognitive processes (e.g., Hackley & Valle-Inclán, 1998; Miller & Hackley, 1992; Osman, Bashore, Coles, Donchin, & Meyer, 1992; Sangals, Sommer, & Leuthold, 2002; Ulrich, Leuthold, & Sommer, 1998; Ulrich, Rinkenauer, & Miller, 1998). Generally, prior to a motor response a slow negativity, the readiness potential (Kornhuber & Deecke, 1965), develops over motor cortex. The readiness potential is more pronounced contralateral than ipsilateral to the responding hand (Vaughan, Costa, & Ritter, 1968). To quantify this asymmetry, the LRP is obtained by subtracting activation over the primary motor cortex ipsilateral to the response side from contralateral activation, and then averaging the difference waveforms across hands (for a detailed description, see Coles, 1989). By that means, the development of motor activation can be monitored during some hundred milliseconds preceding a voluntary movement of a single hand.

The LRP onset provides a time marker dividing RT into two intervals, one before and one after the beginning of hand-specific lateralized motor activation. In chronopsychophysiological research, the LRP is used to determine whether premotor (stimulus-locked LRP, S-LRP) or motor processes (response-locked LRP, R-LRP) or both contribute to RT effects caused by certain experimental manipulations. For example, premotor and motor sources of RT decrement/increment found in flanker tasks (Gratton, Coles, Sirevaag, Eriksen, & Donchin, 1988) or precue tasks (Leuthold, Sommer, & Ulrich, 1996) have been investigated with S-LRP and R-LRP.

Clearly, exact estimation of LRP onset represents an important issue. LRP onset is defined as the point in time at which the potential begins to deviate from baseline and becomes negative. In previous studies, various methods of measuring LRP onsets, and differences between LRP onsets, were compared (Miller, Patterson, & Ulrich, 1998; Mordkoff & Gianaros, 2000; Smulders, Kenemans, & Kok, 1996; Ulrich & Miller, 2001). In principle, two issues have to be distinguished with respect to LRP onset detection. The first is scoring techniques (criterion-based, regression-based, and baseline-deviation), and the second is methods of LRP onset estimation (single-subject and jackknifing procedures, respectively; see below). Regarding the first distinction, criterion-based scoring techniques determine the LRP onset as the first point in time that the LRP exceeds a predefined, arbitrary criterion. This criterion can be defined in terms of absolute values (e.g., 0.5 μV or 1.0 μV), or as a certain proportion of the maximum LRP amplitude (e.g., 50%). With the regression-based scoring method, the LRP onset is determined as the point of intersection of two regression lines. The first regression line fits the baseline before LRP onset, whereas the second line models the rise of the LRP after the onset. Depending on the specific method, however, several restrictions apply to the two regression

We thank Ines Jentzsch, J. Toby Mordkoff, and Fren Smulders for helpful comments on earlier drafts of the manuscript.

Address reprint requests to: Jutta Stahl, Georg Elias Mueller Institute for Psychology, University of Goettingen, Gosslerstr.14, D-37073 Goettingen, Germany; e-mail: jstahl@uni-goettingen.de.

lines. For example, with the *1-df* regression-based scoring technique, the "pre-onset" line is forced to have both zero amplitude and slope. Moreover, the "post-onset" line is required to end at the time and height of the LRP peak. Thus, there is only one degree of freedom, resulting from variations in slope of the post-onset line (for a detailed description, see Schwarzenau, Falkenstein, Hoormann, & Hohnsbein, 1998). Finally, baseline-deviation methods determine the LRP onset as the first point in time that the LRP consistently differs from baseline by a multiple of the standard deviation of baseline (see Osman et al., 1992).

The second distinction relates to two different methods of LRP onset estimation, that is, performing onset detection on single-subject or jackknifed LRP waveforms. Single-subject estimation means that individual LRP onsets are determined from the averaged LRP waveform of each single participant by one of the above-mentioned scoring techniques. The main problem is that individual averaged LRPs are characterized by low signal-to-noise ratio and LRP onset scores could be far away from true LRP onsets. To solve this problem, based on their computer simulations, Miller et al. (1998) recommended the usage of jackknifing procedures (Efron, 1981; Miller, 1974; Quenouille, 1949) in LRP onset estimation. The idea behind jackknifing is to reduce the noise *before* LRP onset detection. In principle, the signal-to-noise ratio is increased by using LRP waveforms averaged across different participants instead of single-subject LRP waveforms. To obtain the jackknifed mean LRP onset score j_i for each participant i ($i = 1 \ldots n$), first, n grand-average waveforms are calculated across participants by successively omitting every participant once. Then, for each of the n grand-average waveforms, the LRP onset is detected with one of the scoring techniques described above. This results in n jackknifed LRP onset scores ($j_1 \ldots j_n$), with each j_i being based on the data from all participants but i. Miller et al. demonstrated that jackknifing is superior to single-subject procedures in detecting a simulated difference in LRP onset between two experimental conditions. However, there is some debate as to whether this superiority of jackknifing critically depends on the scoring technique used. Ulrich and Miller (2001) suggested that jackknifing be applied irrespective of the scoring technique because, in their data, it always performed better than single-subject procedures. By contrast, Mordkoff and Gianaros (2000) have found no consistent superiority of jackknifing and, for example, recommended the single-subject *1-df* regression-based technique for the analysis of LRP data.

Limitation of the Application of Jackknifed LRP Onsets to Factorial Designs

Until now, researchers who wished to analyze jackknife-based LRP onsets had to employ factorial designs. With two experimental conditions, jackknifed LRP onset scores are determined for each participant and condition by assigning a jackknifed score j_i to each participant. This score is defined as the onset of the grand-average LRP of the remaining $n - 1$ participants in that experimental condition (see above). Then, one has to compare the empirically found difference between jackknifed LRP onsets in the two conditions against an appropriate estimation of the standard deviation of jackknifed onset differences, assuming a true zero difference between conditions. Miller et al. (1998) have established a simple solution to this problem, and Ulrich and Miller (2001) extended the applicability of jackknifed LRP onset scores to designs involving more than two conditions, both within subject, and/or between subjects.

Note that in factorial designs no interpretation of individual jackknifed LRP onset scores is being made, as conclusions are based on differences between groups or conditions. For many research questions, however, exactly the information inherent in individual LRP onsets would be highly welcome. For example, one may hypothesize that intelligence is related to speed of sensory information processing, which would make it desirable to correlate some intelligence measure with individual S-LRP onset latency, rather than latency of the overt response, which also reflects speed of motor processes. Yet, as noted above, single-subject LRP onsets may not be very reliable because they are rather strongly affected by EEG noise (except, perhaps, with regression-based procedures). On the other hand, jackknifing would substantially increase reliability of LRP onset detection, but it is not known how a correlation between individual LRP onsets and some external variable like intelligence would change when jackknifed LRP onsets were used. These two reasons may be partly responsible for the fact that, until now, the LRP has only rarely been employed in personality research (Posthuma, Mulder, Boomsma, & de Geus, 2002; Rammsayer & Stahl, 2004).

One may argue that questions regarding interindividual differences in jackknifed LRP onsets may be satisfactorily addressed by employing the median split technique, dividing participants into one group scoring low on the relevant construct, and another group scoring high. Accordingly, Rammsayer and Stahl (2004), in order to test group differences in S-LRP and R-LRP onset between extraverts and introverts, had to split their sample at the median score of Extraversion. Doing so, however, may have caused substantial loss of information about both LRP-onset and Extraversion variability within the two groups, and, strictly, an interpretation in terms of a relationship between LRP onset and Extraversion was not permitted. This is illustrated by the following two examples. First, assume, in a given data set, a true difference in LRP onsets between subjects above and below a certain Extraversion score. However, if this critical score differs from the sample median, analysis involving a factor Extraversion based on median split may fail to detect an LRP onset difference between extraverts and introverts. Consider a second data set where a significant difference in LRP onset between groups was established. Yet, this difference was caused by participants close to the median whereas there was no difference in LRP onset between participants scoring *very* high and low on Extraversion. Then, an interpretation in terms of a linear relationship between LRP onset and Extraversion would be seriously compromised. In either case, if one had the opportunity to estimate correlation coefficients between individual LRP onsets and personality scores, the combined information from correlation and between-group analyses should be superior to between-group analyses alone.

Application of Jackknife-Based LRP Onsets in Correlative Approaches

Given the above considerations, the present study had two main goals. First, the mathematical basis for the application of jackknife-based LRP onsets to correlative approaches should be provided. In a second step, then, several combinations of scoring technique and estimation method (single-subject and jackknifed) should be compared with respect to their utility for the detection of LRP onsets in correlative approaches.

Regarding the first goal, the reader is invited to participate in the following thought experiment. Imagine a grand-average LRP computed by averaging across n individual averaged LRPs that were obtained from n participants. Now, the averaged LRP of participant i is omitted from the grand average. If this leads to an increase in the onset of the grand-average LRP computed for the remaining $n-1$ participants, the individual LRP onset of participant i must have been rather *early*. However, with jackknifing, he/she is assigned an onset that is rather *late*, because it is determined from the grand-average LRP of $n-1$ participants except him-/herself. Thus, the distribution of jackknifed LRP onsets in a group represents something like a mirror image (or, in mathematical terms, the "inverse") of the distribution of the individual, single-subject LRP onsets.

Now, consider correlations. Imagine, first, n true *original* LRP onsets o_i for each individual in an empirical sample, and a perfectly reliable measurement for these onsets. Furthermore, within our sample assume a known relationship r_{ox} between true individual LRP onsets o_i and values x_i of an external variable. With given sets of o_i and x_i, jackknifed LRP onset scores j_i are fixed, as well as their correlation with the external variable, r_{jx}. Under these circumstances, there must be a mathematical expression describing the relationship between the two correlation coefficients r_{jx} and r_{ox}. Indeed, after performing a few algebraic operations we could prove that r_{jx} can quite easily be derived from r_{ox}. Surprisingly, r_{jx} simply equals -1 times r_{ox} (see Appendix, Proof A):

$$r_{jx} = -r_{ox}. \tag{1}$$

Therefore, to compute correlation coefficients from jackknifed LRP data, the following procedure is suggested. First, an averaged individual LRP waveform is calculated for each of the n participants. Second, from these waveforms, n grand averages are computed for all participants but i, and n jackknifed LRP onsets are determined according to one of the above-mentioned scoring techniques. Then, the correlation coefficient between these jackknifed LRP onsets and the variable of interest is computed. After reversing the sign, the resulting coefficient provides an estimate for the true correlation between individual LRP onsets and values on the external variable.

So far, we have shown that the correlation between jackknifed LRP onsets and an external variable can serve as an estimate of the correlation between the original, individual LRP onsets, and the external variable. To demonstrate the utility of jackknifed LRP onsets in correlative approaches, in the next step we provide evidence for the presumed superiority of jackknife-based over single-subject correlation coefficients. To this end, simulation experiments are run, comparing several combinations of estimation methods (single-subject and jackknifed) and scoring techniques (absolute and relative criteria and a regression-based technique) in their ability to detect a true correlation introduced into the data. Because selection of an appropriate simulation procedure represents a highly important issue, in the following we briefly discuss simulation procedures that have been

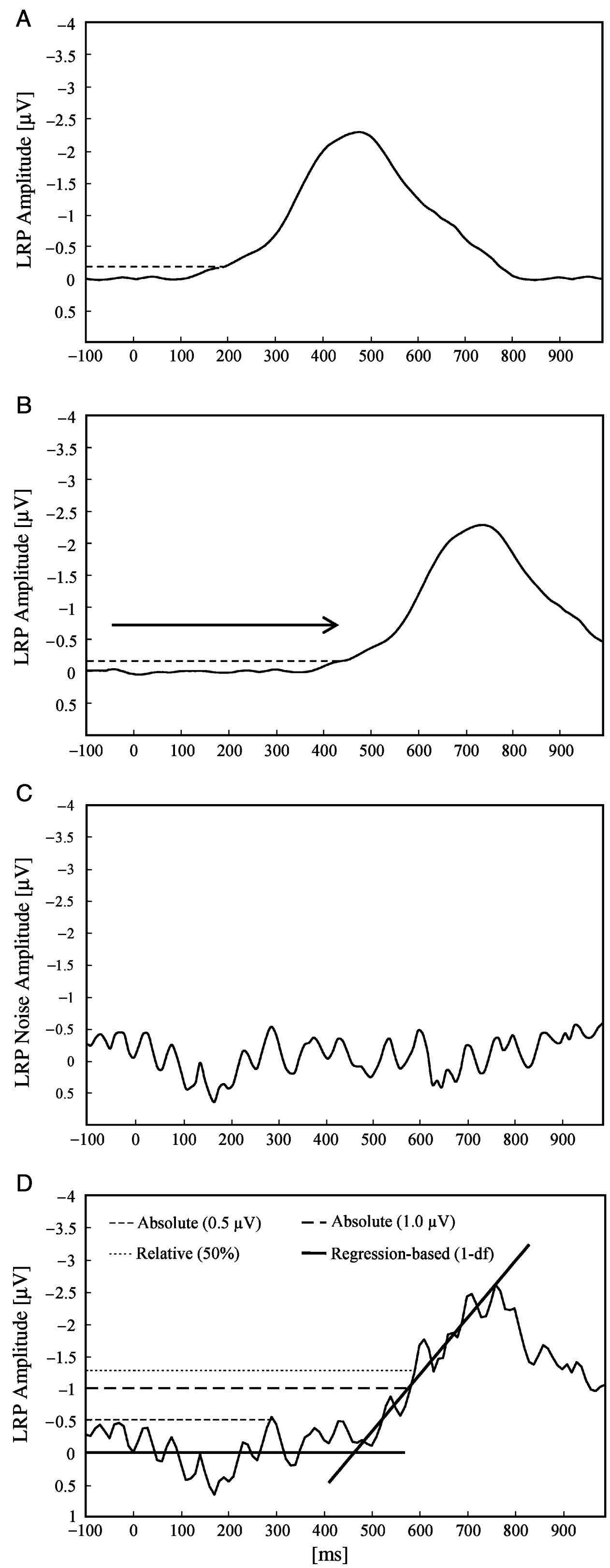

Figure 1. Schematic illustration of the procedure used to obtain realistic LRP waveforms with known true onsets. A: A smooth, individual averaged LRP is drawn from the pool of 100 true waveforms, and the onset is determined according to the 0.2-μV criterion. B: The waveform is shifted in time by the amount required for true LRP onset to match a value that gives rise to a predefined correlation with an external variable; for details, see text. C: An LRP noise waveform is drawn from a pool of 1,000 waveforms and rescaled to match a desired noise intensity of either 0.6 μV, 1.0 μV, or 1.4 μV. D: The rescaled noise waveform is added to the shifted LRP waveform. After the procedure has been performed n times to obtain the n realistic LRP waveforms of one simulated experiment, onset detection is performed according to one of the scoring techniques (0.5-μV, 1.0-μV absolute criterion, 50%-relative criterion, and regression based technique), once with the single-subject and once with the jackknifing procedure.

employed earlier in the literature. Thereafter, a novel simulation procedure more appropriate for the present approach is put forward.

Simulation Methods Applied in Previous Studies

In two simulation studies (Miller et al., 1998; Mordkoff & Gianaros, 2000, first set of simulations), real EEG data were used to evaluate different methods of LRP onset detection. For example, to simulate a 100-ms experimental effect on S-LRP onset, single-trial EEG waveforms recorded over left and right motor cortex were shifted forward in time by 100 ms in the experimental condition, but left unchanged in a control condition. The same procedure applied to RT. Using this technique, Miller et al. found an effect in jackknifed S-LRP but not R-LRP onset, as was expected. Shifting original EEG data in time creates highly realistic data, and the magnitude of the effect can be precisely controlled. Interindividual and intertrial variability of LRP onset, EEG noise, and RT can be assumed externally valid.

Another way to simulate LRP data was described by Smulders et al. (1996) and Mordkoff and Gianaros (2000, second set of simulations). Briefly, in both studies background EEG was created by means of a second-order autoregression procedure, adding a normally distributed noise parameter to each data point. The LRP was then added to the background EEG as the ascending part of a sine wave with variable amplitude and period parameters (for details, see Smulders et al., 1996). With this procedure, noise intensity as well as LRP slope were known and could be varied, which represents the advantage of artificial LRP waveforms. The authors showed that accuracy of onset detection varied with both LRP slope and EEG noise.

Simulation Method Applied in the Present Study

Both these simulation procedures described were less suitable for the present correlative approach. Employing only synthetic LRP samples was not appropriate because our present study, which focuses on interindividual differences, has to ensure realistic, that is, representative, between-subjects variability in LRP slope and magnitude. Therefore, it appeared most reasonable to employ variability in the LRP waveform as it occurs *empirically*. On the other hand, as opposed to Miller et al. (1998), within our correlative approach the idea of shifting original, noisy single-trial EEG waveforms in time could not be applied. This is because our approach does not only involve *differences* between LRP onsets disregarding onsets themselves, but requires assessing true, predefined individual LRP onsets. Only then can true predefined correlations between LRP onset and an external variable be introduced into the data.

In the following, the advantages of both earlier simulation procedures are utilized, and an amalgamation of both is suggested. The new simulation method, on the one hand, allows employing real human EEG data, ensuring external validity. On the other hand, several parameters such as magnitude of the predefined correlation, LRP noise intensity, and between-subjects variability in true LRP onset can be exactly controlled. The simulation procedure run during each simulated experiment involved three steps described in the following (see Figure 1 and Table 1): (1) drawing "true" LRPs[1] from a pool of very low-noise waveforms varying in amplitude and shape, (2) introducing a predefined correlation coefficient between true LRP onsets and the external variable, by shifting true waveforms in time, and (3) adding noise of different intensity to the shifted true LRP waveforms. Finally, single-subject and jackknifed LRP onsets were determined by means of different scoring techniques (see below), and correlations between the obtained LRP onset estimates and values of the external variable were recorded.

Table 1. *Example for Obtaining a Predefined Correlation r Between an External Variable (x_i) and True LRP Onsets (o_i)*[a]

	LRP onset	
External variable (x_i)	Initial array	Final array (o_i)
200	200	200
250	250	350
300	300	250
350	350	300

[a]The correlation *r* with the external variable equals 1.0 for the initial LRP onset array. The initial array is permuted to create a final array with the desired correlation of 0.4.

1. Creating true LRP waveforms. The pool of true LRP waveforms contained 100 very low-noise LRPs, each representing the averaged LRP of one of 100 participants in a real experiment. Each of the individual averaged LRPs was based on artifact-free EEG epochs from 1,000 trials of a choice-RT experiment recently run in our laboratory.[2] The EEG was derived from electrode positions C3′ and C4′, which were located 1 cm anterior to positions C3 and C4 of the 10–20 system (Jasper, 1958), respectively. The recording was continuous; sampling rate was 500 Hz, and time constant was infinite. The EEG was filtered with a 70-Hz low-pass and segmented off-line, with epochs ranging from 300 ms before to 900 ms after onset of the imperative stimulus. Single-subject averaged LRPs were calculated as usual (cf. Coles, 1989) for each of the 100 participants, and then filtered with a low-pass of 10 Hz. Among the 100 resulting true LRP waveforms in the pool, LRP amplitude and slope ranged from 0.6 μV to 6.1 μV and from 5.9 μV/s to 12.4 μV/s, respectively.

2. Producing predefined correlations. Figure 1 illustrates the method of introducing a true correlation between individual LRP onset and an external variable. First, depending on sample size *n* within a set of simulations (40 and 60, see below), *n* true LRP waveforms were randomly drawn without replacement from the pool of 100 true LRP waveform (cf. step 1), separately for each simulated experiment (see Figure 1A).

In the next step, a score x_i on the external variable was assigned to each of the *n* simulated participants. To this end, *n* normally distributed random values were drawn, with a mean of

[1]The term *true LRP waveform* was used to refer to the extremely smooth LRP waveforms from which *true LRP onsets* were determined by means of the 0.2-μV absolute criterion. We are aware of the fact that also the smooth LRP waveforms were not entirely noise free. For reasons of simplicity and clarity, however, in the remainder of this article we refer to true LRP waveforms and true LRP onsets as opposed to LRP waveforms and onsets that were obtained after LRP noise had been added.

[2]True LRP waveforms and LRP-noise waveforms were obtained from a simple visual choice RT experiment. Accessory auditory stimuli were presented *after* the visual imperative stimuli. Ten conditions each involving 100 trials resulted from variations in tone intensity and delay. Left- and right-hand responses, respectively, were required to two different visual stimuli ("v" and "w"). No advance information regarding the correct response was available.

350 ms and a standard deviation of 50 ms (which corresponded to parameters observed in the empirical study from which the LRP data were taken), and written into an array of x-values (see Table 1, first column). Another n-dimensioned array, the *initial LRP onset array* (see Table 1, second column), was then equated to the x-value array. This resulted in a correlation of $r = 1.0$ between initial LRP onsets, and x values. Then, initial LRP onsets were randomly permuted until the required true correlation of, for example, $r_{ox} = .40$ between final LRP onsets o_i and values x_i was achieved (Table 1, second and third columns). The iterative permutation procedure stopped when the required correlation coefficient was approximated at the fifth decimal place. After having obtained two arrays, an x-value array and a *final LRP-onset array*, which correlated with the desired coefficient, the onsets of the n true LRP waveforms drawn earlier (Figure 1A) were determined using a very low absolute criterion of 0.2 μV. Note that this onset detection was indeed extremely reliable, because the pool of true LRP waveforms contained very low-noise, averaged LRP waveforms. Careful visual inspection prior to the simulations had ensured that in any of the true LRPs there was no preactivation greater than 0.2 μV. Thus, the 0.2-μV criterion always produced true LRP onsets appropriately reflecting the point in time where, obviously, the LRP began to deviate from zero. Finally, each true LRP waveform was shifted by exactly the amount of time that was required for its onset to match the value given by the final LRP-onset array (Figure 1B).

3. Adding LRP noise. At this point, there is a set of n LRP waveforms with defined true onsets (according to the 0.2-μV criterion), correlating to a predefined coefficient with an external variable. Yet, the simulations required realistic, noisy data. For this purpose, there was a second pool of data containing averaged-LRP noise that was extracted from data of the same 100 subjects. To ensure that noise waveforms were not contaminated by event-related potentials, artifact-free baseline segments were obtained from interstimulus intervals (ISIs) of the above-mentioned experiment. Each ISI was computed as the sum of a constant period of time of 1,000 ms and an exponentially distributed random variable with a mean of 2,000 ms, which served to minimize anticipatory responses. To approximate realistic noise intensity at the level of single-subject averaged LRPs, 10 noise waveforms were calculated per subject, each being based on 100 different trials (50 per responding hand). LRP noise was computed like the LRP itself, by computing averaged $C3' - C4'$ and $C4' - C3'$ difference waveforms if the following imperative stimulus required a right-hand and a left-hand response, respectively, and then averaging across hands (Coles, 1989). Thereby, potential slow-wave activity due to the participant's preparation for the imperative stimulus was excluded from LRP noise segments (left/right responses to the imperative stimulus could not be predicted in advance). This resulted in a pool of 1,000 LRP noise waveforms. Across all noise waveforms, mean amplitude was normally distributed, with a grand mean of 0.0 μV. Mean standard deviation within one noise waveform was 0.9 μV.

After a predefined correlation between LRP onset and the external variable had been achieved during each simulated experiment (see *Producing Predefined Correlations*), n LRP noise segments were randomly drawn without replacement from the pool of 1,000 segments (Figure 1C). Next, all n segments were rescaled to match an LRP noise intensity (i.e., within-segment standard deviation) of either 0.6 μV (low), 1.0 μV (medium), or 1.4 μV (high). The three noise levels were chosen to cover a realistic range of LRP noise intensity at the level of single-subject averages. The obtained n noise segments were then added to the n true LRP waveforms, similar to Smulders et al. (1996). If the estimation method was single-subject, onset detection was performed on the n resulting noisy LRP waveforms (see Figure 1D). Several scoring techniques appropriate for S-LRPs were employed, that is, two absolute criteria (0.5 μV and 1.0 μV), a 50%-relative criterion, and the *1-df* regression-based method (see Schwarzenau et al., 1998). If the estimation method was jackknifing, the same scoring techniques applied to the n grand-average LRPs that were obtained from averaging across $n - 1$ individual LRPs (see above). Finally, both the jackknifed and the single-subject LRP onsets were correlated with the values of the external variable, providing a jackknifed and single-subject estimation, respectively, for the true correlation coefficient. Thereafter, the sets of true LRPs and LRP noise waveforms were initialized, and the next simulated experiment was run.

Independent and Dependent Variables

In the present study, several sets of simulations were run involving different onset scoring techniques (0.5-μV criterion; 1.0-μV criterion; 50% of maximum LRP amplitude; and *1-df* regression-based), estimation methods (single-subject and jackknifed), sample sizes ($n = 40$ and $n = 60$), predefined true correlations (0.0 to 1.0, in steps of 0.1), and LRP noise intensities (low, medium, and high). The 50%-relative criterion and the *1-df* regression-based method were chosen because of the recommendations by Miller et al. (1998) and Mordkoff and Gianaros (2000), respectively. Two absolute criteria (one low and one high) were employed because, in the literature, no particular criterion has been favored.

Dependent variables were means and standard deviations of estimated correlation coefficients. Moreover, as a measure of statistical power, the proportion of significant estimated correlation coefficients is reported. For this purpose, p was set to .05 (two-tailed), but only positive-significant coefficients in case of single-subject procedures and negative-significant correlations in case of jackknifing were considered, given a simulated true correlation greater than zero. If the true correlation was zero, proportion of significant ($p < .05$) coefficients was computed disregarding their direction, providing a measure of Type I error. For each estimation method, onset scoring technique, sample size, noise intensity, and each of the 11 predefined true correlations between LRP onset and the external variable (0.0 – 1.0), the simulation experiment described above was run 1,000 times.

Results

Mean Estimated Correlation Coefficients

Figures 2 and 3 depict means and standard deviations of estimated correlation coefficients as a function of estimation method (single-subject and jackknifing), scoring technique (0.5-μV criterion; 1.0-μV criterion; 50% criterion, and *1-df* regression-based), predefined correlation coefficient (0.0 – 1.0), and noise intensity (low, medium, and high; panels A–C), for sample sizes of $n = 40$ (Figure 2) and $n = 60$ (Figure 3).

As can be seen, with both estimation methods, all four onset criteria and all three noise intensities mean estimated correlation starts at 0.0 for a predefined coefficient of 0.0, and then linearly increases with increasing predefined coefficient. However, steepness of this linear increase and, hence, the mean estimated correlation coefficient that is obtained with a predefined coefficient

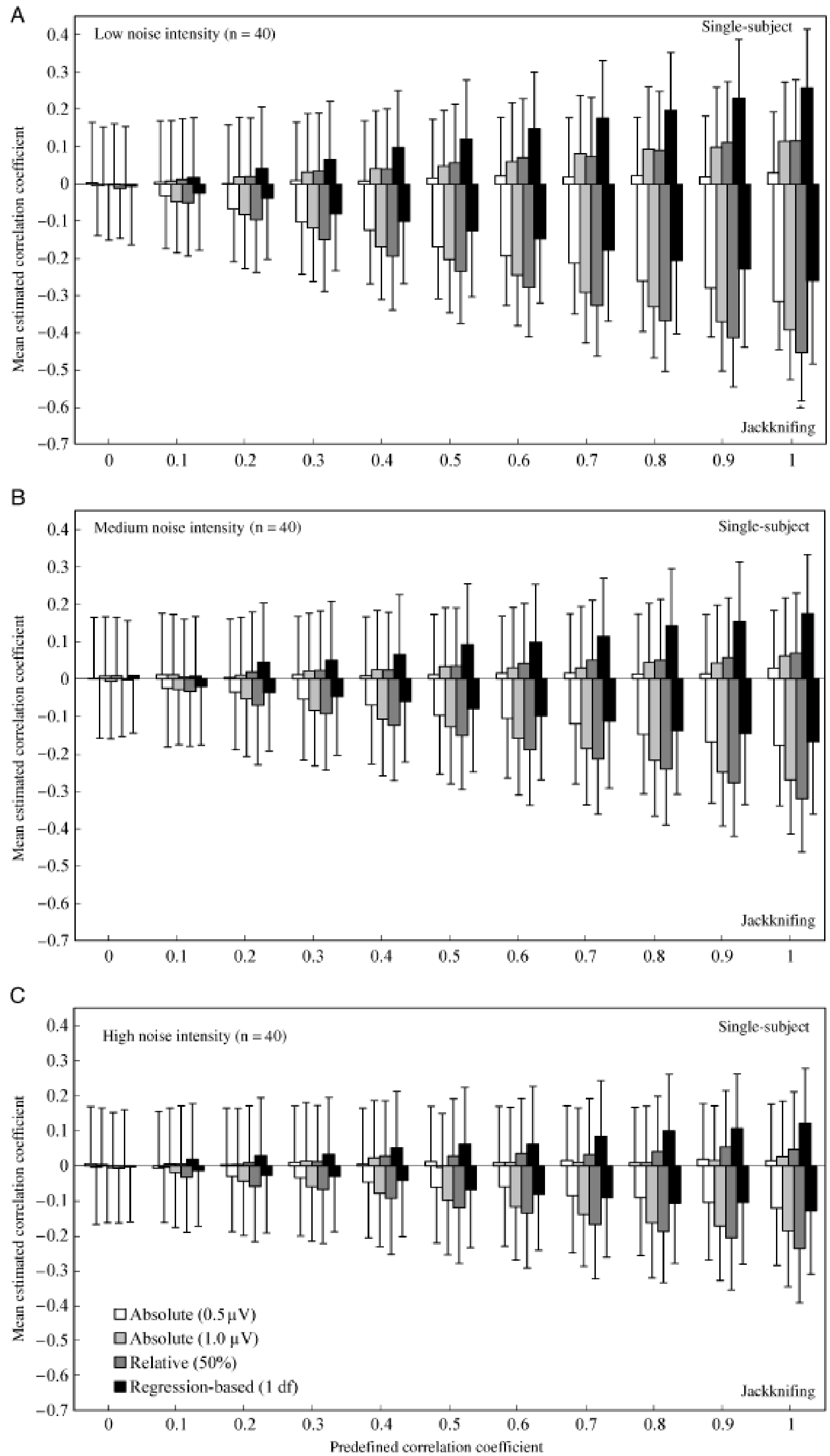

Figure 2. Mean estimated correlation coefficient and standard deviation as a function of scoring technique, predefined correlation coefficient, estimation procedure (single-subject and jackknifing), and LRP-noise intensity (A: low, B: medium, and C: high). Sample size $n = 40$.

of 1.0, markedly varies between conditions. With other conditions held constant, for all scoring techniques except the *1-df* regression-based, mean jackknifed correlation coefficients are, by absolute value, much higher than the corresponding mean single-subject coefficients. Depending on onset criterion, jackknifing performs up to 10 times better than the single-subject

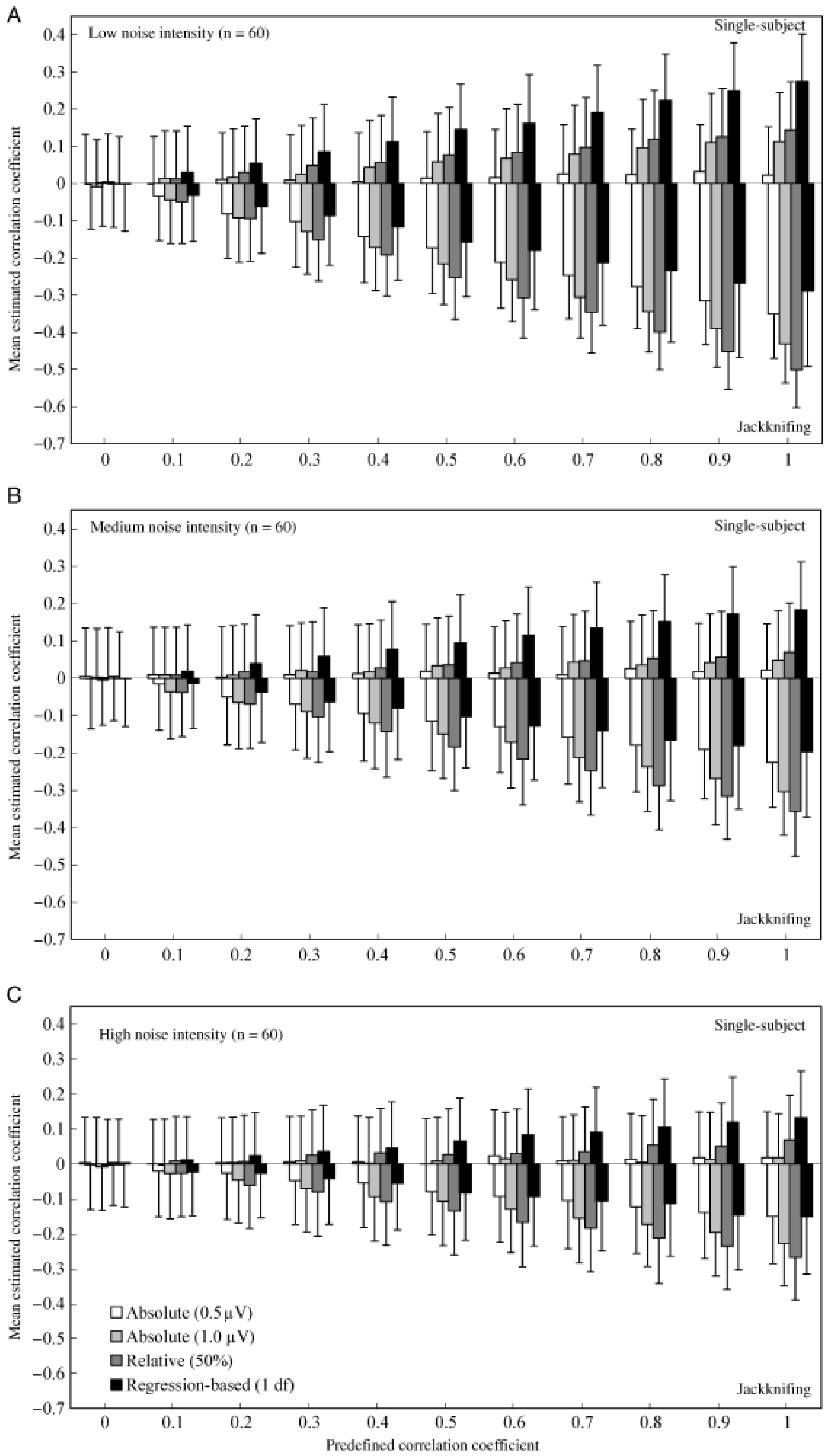

Figure 3. Mean estimated correlation coefficient and standard deviation as a function of scoring technique, predefined correlation coefficient, estimation procedure (single-subject and jackknifing), and LRP-noise intensity (A: low, B: medium, and C: high). Sample size $n = 60$.

method. For example, the 0.5-μV criterion produces, with low noise, $n = 40$, and predefined $r = 1.0$, a mean single-subject coefficient of $r = .03$ and a mean jackknifed coefficient of $r = .32$ (see Figure 2A). Note that, as a result of unreliability, mean estimated correlation coefficients can only be lower than the given true correlation, never higher. Therefore, the height of the mean estimated correlation coefficient observed for a given true correlation and noise intensity directly indexes performance of an LRP onset detection method.

Higher noise intensity markedly reduces mean estimated correlation coefficients. Whereas, for example, with $n = 40$, low noise, and a true correlation of 1.0 a mean estimated correlation of .45 is obtained with a combination of jackknifing and a 50% criterion, the same procedure yields only a mean coefficient of .24 with the high noise intensity (see Figure 2A,C). Note that single-subject procedures are at least as much affected by increases in LRP noise. The variation in sample size, on the other hand, has rather little effect on mean estimated correlation coefficients. For example, coefficients obtained with a true r of 1.0 and the 50%-jackknife method are about the same with $n = 40$ and $n = 60$ ($r = .45$ and .50, respectively; see Figures 2A and 3A).

Remarkably, with the *1-df* regression-based technique jackknifing does not perform better than the single-subject procedure. Indeed, when considering only single-subject coefficients that are generally low, the regression-based method yields by far the highest mean estimated correlation coefficient, which is $r = .28$ ($n = 60$; predefined $r = 1.0$; low noise). Note, however, that, for a given sample size and noise intensity, the 50%-jackknife procedure always produces the highest mean estimated correlation coefficient. Depending on noise intensity, also mean jackknifed coefficients obtained with a 0.5-μV criterion and/or 1.0-μV criterion are higher than those based on the *1-df* regression procedure, either jackknifed or single-subject. For example, with medium noise and $n = 60$ both the single-subject and the jackknifed regression-based procedure estimate an averaged r of .19 if true correlation was 1.0, whereas jackknifing with the 0.5-μV criterion and the 1.0-μV criterion yields mean rs of .23 and .31, respectively (see Figure 3B).

Regarding the standard deviation of estimated coefficients, with all scoring techniques but the regression-based there is a slight advantage of jackknifing over single-subject procedures (e.g., .11 and .13, respectively, with $n = 60$ and medium noise; see Figure 3B). The regression-based technique, however, does not benefit from jackknifing also in terms of variability of the estimates (standard deviations of 0.12 and 0.17 with single-subject and jackknifing procedures, respectively; see Figure 3B).

This variability advantage of jackknifing found with most scoring techniques is relevant to the empirical praxis, because to

Table 2. *Proportion of Significant Estimated Correlation Coefficients ($p < .05$, Two-Tailed) as a Function of Predefined Correlation Coefficient (r), EEG-Noise Intensity, Scoring Technique (Absolute Criterion: 0.5 μV and 1.0 μV, Relative Criterion 50%, and Regression-Based [RB]), and Estimation Method (Jackknifing and Single-Subject)*[a]

	Single-subject				Jackknifing			
Noise intensity and predefined r	0.5 μV	1.0 μV	50%	RB	0.5 μV	1.0 μV	50%	RB
Low noise intensity								
0.0	.059	.041	.052	.047	.042	.041	.029	.044
0.1	.052	.064	.062	.063	.070	.045	.076	.066
0.2	.051	.064	.064	.091	.103	.103	.122	.093
0.3	.060	.078	.075	.118	.145	.163	.210	.111
0.4	.050	.096	.091	.141	.162	.267	.309	.168
0.5	.055	.077	.098	.178	.266	.362	.424	.241
0.6	.058	.102	.111	.234	.330	.450	.549	.277
0.7	.059	.134	.118	.304	.382	.575	.688	.362
0.8	.061	.175	.137	.345	.511	.717	.791	.432
0.9	.054	.154	.189	.438	.560	.792	.870	.489
1.0	.068	.186	.198	.512	.640	.831	.910	.573
Medium noise intensity								
0.0	.058	.053	.048	.046	.047	.057	.048	.039
0.1	.069	.068	.044	.060	.074	.055	.063	.064
0.2	.044	.040	.061	.083	.071	.086	.127	.079
0.3	.054	.061	.066	.077	.107	.120	.138	.081
0.4	.051	.066	.058	.107	.117	.163	.181	.110
0.5	.067	.071	.075	.156	.173	.188	.218	.139
0.6	.047	.076	.084	.147	.177	.259	.333	.185
0.7	.055	.076	.088	.165	.196	.326	.390	.217
0.8	.056	.084	.091	.228	.259	.388	.467	.257
0.9	.057	.073	.101	.240	.288	.496	.566	.289
1.0	.061	.102	.120	.303	.316	.537	.683	.350
High noise intensity								
0.0	.046	.051	.043	.054	.069	.060	.055	.049
0.1	.048	.059	.063	.062	.048	.077	.065	.058
0.2	.052	.049	.059	.083	.070	.074	.110	.080
0.3	.061	.067	.063	.064	.083	.095	.109	.079
0.4	.054	.075	.077	.094	.097	.113	.158	.082
0.5	.051	.036	.070	.105	.087	.147	.187	.132
0.6	.062	.047	.066	.118	.125	.186	.216	.126
0.7	.056	.058	.079	.142	.141	.207	.276	.165
0.8	.050	.056	.088	.161	.147	.256	.321	.188
0.9	.066	.056	.095	.156	.179	.280	.382	.196
1.0	.056	.074	.093	.190	.194	.334	.468	.237

[a]Data are based on 1,000 simulations and a sample size of $n = 40$.

obtain a significant correlation is often more important than the numerical value of the coefficient itself. However, the reader may have noticed the generally low values of mean estimated correlation coefficients reported so far. For example, for a predefined r of 1.0 mean estimated correlation coefficient does not exceed a value of $r = .50$ even with jackknifing, lowest noise, and largest sample size (see Figure 3A). Without jackknifing, maximal r was only .28 (regression-based, low noise; see Figure 3A). In a single empirical study, results like this may often fail to reach significance, because sample sizes fairly below 100 are typical of electrophysiological research on personality (e.g., Coyle, Gordon, Howson, & Meares, 1991; Stelmack, Houlihan, & McGarry-Roberts, 1993; Walhovd & Fjell, 2002). Consequently, with reasonable sample size and EEG noise intensity, even a perfect correlation ($r = 1.0$) between true LRP onset and some external variable may often fail to be detected in the data when single-subject onset detection is used. Thus, besides estimating numerically greater correlation coefficients, another empirically relevant advantage of jackknifing may be the greater capability to detect, in one single experiment, an existing true correlation different from zero between LRP onset and an external variable.

Proportion of Significant Estimated Correlation Coefficients

To further investigate this assumed advantage of jackknifing, in Tables 2 and 3 the proportion of significant estimated correlation coefficients ($p = .05$; see *Independent and Dependent Variables*), providing a measure of statistical power, is shown as a function of predefined correlation coefficient, jackknifed versus single-subject estimation, scoring technique, and noise intensity, separately for $n = 40$ (Table 2) and $n = 60$ (Table 3).

With a predefined correlation of 0.0 and medium noise intensity, both the single-subject procedure and jackknifing only rarely obtain significant correlation coefficients (probability around 5%, two-tailed), both with $n = 40$ and $n = 60$. Hence, there is no difference between estimation methods with respect to Type I error. Increasing true correlation causes a relatively rapid increase in number of significant correlation coefficients when onset estimation is jackknife-based or the single-subject *1-df* regression-based method is employed. Again, the combination of jackknifing with a 50%-relative criterion performs best, as is evident, for example, in a more than 99% probability for a true 1.0 correlation to be detected in the data (see Table 3, $n = 60$, low noise). The single-subject regression-based method again does

Table 3. *Proportion of Significant Estimated Correlation Coefficients ($p < .05$, Two-Tailed) as a Function of Predefined Correlation Coefficient (r), EEG-Noise Intensity, Scoring Technique (Absolute Criterion: 0.5 μV and 1.0 μV, Relative Criterion 50%, and regression-based [RB]), and Estimation Method (Jackknifing and Single-Subject)*[a]

	Single-subject				Jackknifing			
Noise intensity and predefined r	0.5 μV	1.0 μV	50%	RB	0.5 μV	1.0 μV	50%	RB
Low noise intensity								
0.0	.052	.040	.052	.054	.041	.033	.039	.047
0.1	.041	.051	.058	.067	.069	.083	.072	.072
0.2	.043	.063	.067	.099	.133	.146	.131	.114
0.3	.047	.059	.103	.161	.193	.237	.284	.170
0.4	.063	.095	.115	.196	.290	.365	.443	.269
0.5	.052	.119	.151	.298	.374	.524	.649	.374
0.6	.061	.133	.153	.359	.492	.675	.803	.464
0.7	.086	.149	.187	.427	.613	.799	.882	.549
0.8	.059	.182	.241	.550	.722	.890	.952	.613
0.9	.075	.226	.259	.608	.800	.950	.988	.684
1.0	.075	.215	.293	.717	.870	.977	.999	.708
Medium noise intensity								
0.0	.046	.047	.056	.043	.048	.046	.038	.047
0.1	.047	.052	.055	.055	.047	.086	.064	.052
0.2	.055	.062	.062	.087	.101	.113	.113	.088
0.3	.060	.063	.070	.124	.121	.154	.178	.126
0.4	.055	.058	.074	.145	.173	.230	.278	.152
0.5	.070	.074	.083	.173	.222	.307	.409	.216
0.6	.057	.073	.103	.224	.256	.382	.534	.292
0.7	.055	.096	.104	.254	.335	.519	.618	.333
0.8	.064	.095	.097	.310	.406	.590	.745	.406
0.9	.059	.098	.096	.395	.446	.678	.820	.459
1.0	.060	.117	.151	.416	.536	.773	.879	.517
High noise intensity								
0.0	.049	.050	.043	.045	.047	.049	.032	.042
0.1	.042	.053	.047	.054	.076	.075	.075	.067
0.2	.060	.046	.063	.063	.074	.097	.107	.066
0.3	.049	.057	.075	.084	.102	.118	.154	.097
0.4	.051	.048	.083	.092	.096	.174	.205	.116
0.5	.048	.055	.076	.121	.136	.209	.268	.155
0.6	.072	.068	.084	.152	.174	.247	.380	.201
0.7	.048	.057	.076	.167	.223	.338	.414	.224
0.8	.063	.053	.107	.230	.253	.369	.514	.260
0.9	.065	.064	.082	.230	.282	.461	.578	.353
1.0	.071	.064	.119	.272	.332	.556	.670	.376

[a]Data are based on 1,000 simulations and a sample size of $n = 60$.

remarkably better than any other single-subject procedure, but never reaches performance of (nonregression-based) jackknifing procedures. Specifically, unlike with mean estimated coefficients, jackknifing combined with the 0.5-μV criterion also consistently produces a higher proportion of significant correlations than the single-subject *1-df* regression-based procedure. This pattern of results is most likely due to the above-mentioned variability advantage of jackknifing. Consequently, even if mean estimated correlation coefficient is not higher, a greater portion of these estimates becomes significant because values *much* below the mean are less frequently observed.

Interindividual LRP Onset Variability and the Smearing Problem

So far, the applicability of jackknifed LRP onsets to correlative approaches has been successfully demonstrated, and a consistent superiority has been shown for a combination of jackknifing with the 1.0-μV absolute criterion and the 50%-relative criterion, respectively, over the most efficient single-subject procedure, that is, the *1-df* regression-based method. However, there may still be doubts regarding the superiority of jackknifed LRP onsets in correlative approaches, at least as long as a critical issue in jackknifing, interindividual LRP onset variability (Miller et al., 1998; Mordkoff & Gianaros, 2000; Smulders et al., 1996), has not been addressed.

On the one hand, the main advantage of jackknifing is to increase the signal-to-noise ratio before onset detection by including more data. On the other hand, the main disadvantage of jackknifing results from the fact that more data is being averaged, and the analyzed signal is more smeared. Because of smearing, LRP onsets determined from waveforms averaged across trials are biased toward the onset of the fastest trial (see Meyer, Osman, Irwin, & Yantis, 1988). As Miller et al. (1998) noted, jackknifing should be particularly sensitive to smearing because LRP onsets are determined from grand averages. Depending on scoring technique, then, jackknifed LRP onsets are more or less biased toward the onset of the fastest participant. Thus, jackknifing tends to underestimate individual LRP onsets, and, obviously, the greater interindividual LRP onset variability, the greater onset underestimation. Regarding the present study, one may ask whether the superiority of jackknifing is tied to the specific (and, perhaps, low) interindividual variability introduced into true LRP onsets, and whether it would disappear with higher

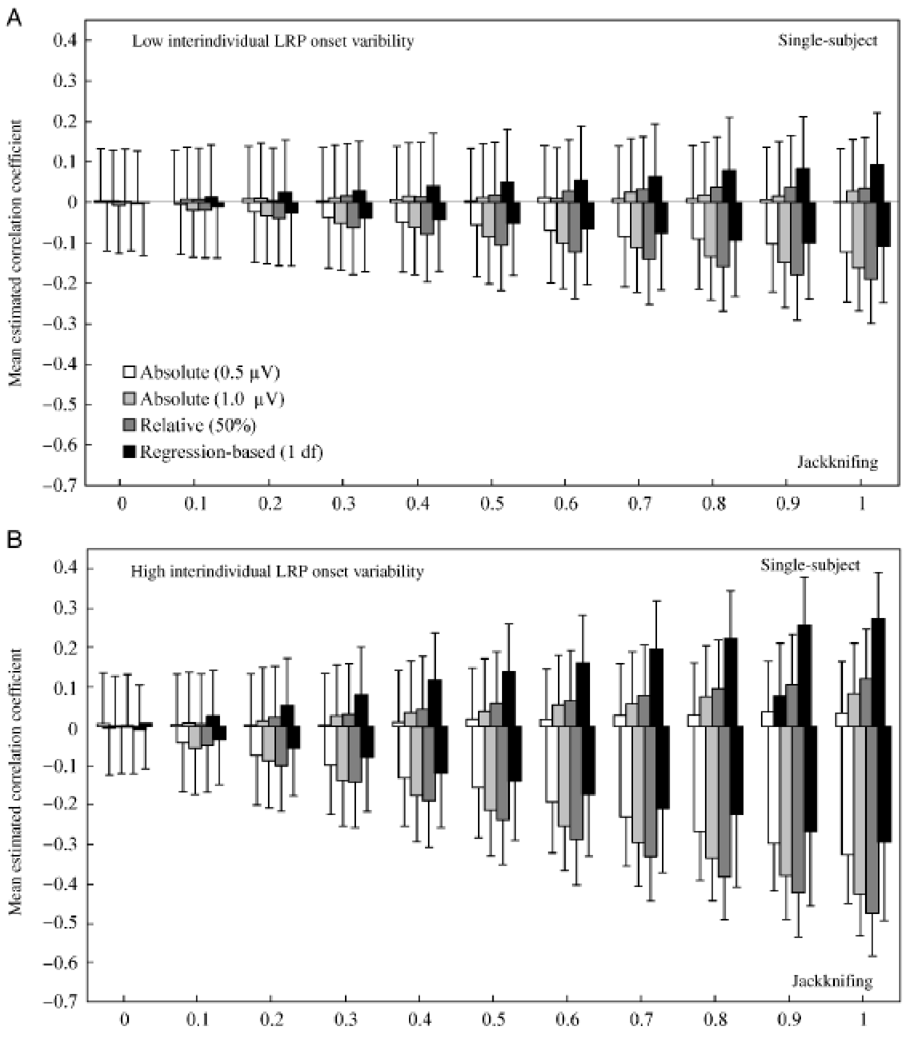

Figure 4. Mean estimated correlation coefficient and standard deviation as a function of scoring technique, predefined correlation coefficient, estimation procedure (single-subject and jackknifing), and interindividual LRP onset variability (A: low and B: high). Sample size $n = 60$; medium LRP-noise intensity.

variability, or stronger smearing. This is especially relevant to the present correlative approach where we are interested in individual LRP onsets themselves, and not merely in differences in LRP onset between two conditions.

As mentioned above (see *Simulation Method*), our simulations were based on a mean true LRP onset latency of 350 ms and a standard deviation of 50 ms, which mimicked the parameters observed in the empirical data employed. From this perspective, the simulations reported so far are externally valid, and the superiority of jackknife-based over single-subject correlations is unlikely to result from interindividual LRP onset variability being too small. Yet, to investigate the effect of interindividual LRP onset variability on the pattern of correlations produced by the different methods, two additional sets of simulations were run, employing lower (i.e., 25 ms) and higher (i.e., 75 ms) true interindividual LRP-onset variability than our prior simulations. The simulation method was the same as above, with the following exception. During the second phase of the simulation procedure (*Creating Predefined Correlations*, see *Simulation Method*), values of the external variable (which then were equated to values in the initial LRP onset array) were drawn as normally distributed random values with means of 350 ms and standard deviations of 25 ms and 75 ms.

Figure 4 shows mean estimated correlation coefficients as a function of estimation method (single-subject and jackknifed), scoring technique, predefined correlation coefficient, and low versus high interindividual variability in true LRP onset. For reasons of clarity, results are reported only for medium noise intensity (within-waveform standard deviation of 1.0 μV) and $n = 60$. Interestingly, mean estimated correlation coefficients were generally much higher with higher interindividual variability in true LRP onsets, or stronger smearing. The pattern of results found earlier for the different combinations of scoring techniques and estimation methods, however, was left unchanged. Specifically, also with high interindividual LRP-onset variability, involving stronger smearing, all scoring techniques but regression produced jackknife-based correlation coefficients that were greater than the maximum single-subject coefficient (see Figure 4B). Again, the combination of jackknifing with a 50%-relative onset criterion yielded the highest estimates. Regarding the proportion of statistically significant estimated correlations, or statistical power, both single-subject and jackknifing procedures benefited from the increase in interindividual LRP onset variability (see Table 4). Although the relative increase in power was stronger for single-subject procedures (e.g., from .166 to .697 with true $r = 1.0$, regression-based), jackknifing showed by far the highest absolute power (.98, with true $r = 1.0$, 50%-relative criterion; see Table 4).

Testing Statistical Significance

May the superiority of jackknifed over single-subject LRP onsets be the result of inappropriate use of statistical tests, as would be the case with noncorrected F values when ANOVAs were carried out on jackknifed data (see Miller et al., 1998; Ulrich & Miller, 2001)? These authors have shown that testing hypotheses regarding jackknifed data requires adjustment of statistical parameters such as t or F values. F values, for example, are inflated because, with jackknifing, variance within groups/conditions is reduced by a factor $(n-1)^2$, whereas variance between groups/conditions remains constant. Hence, F values, which compute as between-group divided by within-group variance artificially increase by a factor $(n-1)^2$, and have to be adjusted accordingly (for details, see Ulrich & Miller, 2001). Regarding the present case of correlations, however, we have shown in the Appendix

Table 4. *Proportion of Significant Correlation Coefficients ($p < .05$, Two-Tailed) as a Function of Predefined Correlation Coefficient (r), Interindividual LRP Onset Variability, Scoring Technique (Absolute Criterion: 0.5 μV and 1.0 μV, Relative Criterion 50%, and Regression-Based [RB]), and Estimation Method (Jackknifing and Single-Subject)*[a]

	Single-subject				Jackknifing			
LRP onset variability and predefined r	0.5 μV	1.0 μV	50%	RB	0.5 μV	1.0 μV	50%	RB
Low interindividual variability								
0.0	.050	.042	.052	.044	.055	.041	.038	.053
0.1	.045	.055	.060	.053	.044	.050	.062	.052
0.2	.067	.062	.061	.080	.071	.080	.079	.081
0.3	.052	.057	.064	.079	.080	.109	.130	.095
0.4	.057	.061	.074	.084	.106	.106	.141	.081
0.5	.059	.063	.075	.099	.113	.163	.207	.102
0.6	.052	.060	.076	.112	.130	.165	.232	.146
0.7	.056	.081	.068	.120	.164	.206	.267	.168
0.8	.058	.071	.093	.161	.187	.268	.335	.201
0.9	.063	.057	.081	.159	.189	.325	.413	.205
1.0	.048	.072	.074	.166	.240	.354	.445	.218
High interindividual variability								
0.0	.052	.041	.050	.017	.039	.031	.036	.021
0.1	.050	.054	.056	.044	.086	.087	.086	.050
0.2	.059	.077	.077	.093	.136	.149	.171	.093
0.3	.051	.061	.084	.128	.173	.259	.263	.151
0.4	.070	.089	.099	.206	.264	.381	.436	.249
0.5	.066	.095	.118	.270	.345	.514	.602	.340
0.6	.060	.102	.125	.334	.441	.641	.734	.422
0.7	.084	.124	.146	.448	.586	.765	.839	.548
0.8	.087	.140	.163	.549	.677	.867	.929	.596
0.9	.086	.153	.208	.645	.744	.924	.952	.691
1.0	.083	.157	.226	.697	.818	.968	.980	.712

[a]Data are based on 1,000 simulations, a sample size of $n = 60$, and medium noise intensity.

(Proof A) that a correlation coefficient obtained by jackknifing is valid (after multiplication by -1). Because the statistical significance of a valid correlation coefficient only depends on n, its statistical significance is valid, too.

We can conclude that significance of jackknife-based correlation coefficients can be tested like correlation coefficients in general. Standard statistical software or tables may be used to determine the significance of correlation coefficients, keeping in mind, however, that the sign of the coefficient has to be reversed.

Discussion

Jackknifing of LRP waveforms before onset detection has been recommended in previous studies, but for use with factorial designs only (Miller et al., 1998; Ulrich & Miller, 2001). The purpose of the present study was to extend the application of jackknifing to correlative approaches. To this end, we have shown that the correlation coefficient between jackknifed values and some external variable x equals -1 times the correlation coefficient between the original values and x, assuming all measures were perfectly reliable. In a second step, several sets of simulations were run on realistic LRP data to demonstrate, for different onset criteria, the superiority of jackknife-based over single-subject estimates of a simulated true correlation. For a given true correlation different from zero, with both absolute criteria (0.5 μV and 1.0 μV) and a relative criterion (50%), jackknifing yielded much higher estimated correlation coefficients between LRP onset and external variable than the respective single-subject procedure, consistently across sample sizes and noise intensities. In addition, compared to the corresponding single-subject procedure, jackknifing produced a greater percentage of "correctly" significant estimated correlation coefficients (i.e., coefficients that would lead the researcher to the same conclusion as the true correlation). For a true correlation of zero, jackknifing only rarely, and not more often than single-subject procedures, erroneously indicated a significant coefficient (Type I error around 5%). The single-subject regression-based procedure performed by far more successfully than any other single-subject procedure, but did not benefit at all from jackknifing. Overall, maximum performance was consistently observed for a combination of jackknifing with the 50%-relative criterion.

These findings are interesting in several respects. First, Mordkoff and Gianaros (2000) also reported a slight superiority of single-subject over jackknifed estimates of a simulated experimental effect when onset detection was regression based. The present results for the correlative approach parallel the prior finding, pointing to an incompatibility of jackknifing with regression-based onset detection. A preliminary explanation may be the following. Different onset criteria are differentially affected by different types of LRP noise. One type may be termed *local noise* and another *global noise*. Local noise is high when there is large variability in LRP amplitude in the time range in which LRP onset is detected, and is caused by high-frequency EEG background noise. Global noise may be introduced by interindividual variability in slope and amplitude of the LRP, and by slow potential shifts (e.g., preactivation). Obviously, low absolute criteria are strongly affected by local noise, making LRP onset measures less reliable, but almost unaffected by post-onset global noise (which is, perhaps, introduced by properties of the LRP waveform), as later waveform is not considered. If, then, local noise is markedly reduced by jackknifing, the low absolute criteria perform quite well. Regression-based onset detection, on the contrary, is much less affected by local noise in the range of true LRP onset, because a greater portion of the waveform is considered, but may be sensitive to global noise occurring in the post-onset LRP waveform. Consequently, at the single-subject level, with relatively strong local noise, regression-based onset detection performs outstandingly well. With jackknifing, however, no additional benefit results for regression-based onset scoring, presumably, because jackknifing reduces the same type of error variance (i.e., local noise) that the regression-based method already reduces at the single-subject level. Furthermore, jackknifing introduces global noise, because the LRP waveform is changed. More than absolute criteria, the regression-based technique, which tries to model the rise of the LRP, will be affected by this post-onset global noise. This may be another reason why, for this method, no advantage results from jackknifing.

Still, the reader may be surprised that it was the 50%-relative criterion that, in combination with jackknifing, consistently across noise and interindividual variability conditions most closely approximated the true correlation. This finding is surprising because, naturally, the 50%-relative criterion most often estimates later LRP onsets than do low absolute criteria such as those used here (0.5 μV and 1.0 μV; mean amplitude across all simulations was 1.4 μV with the 50%-relative criterion) or regression-based methods. In other words, the point where the LRP begins to deviate from zero is more directly assessed by the latter two scoring techniques. Regarding the distinction of local and global noise suggested above, the 50%-relative criterion may be considered a bit of both an absolute and a regression-based criterion. Like the absolute criteria, the 50%-relative criterion is affected by local noise occurring in the time range in which the onset is detected and, consequently, strongly benefits from jackknifing. On the other hand, similar to a regression-based criterion, onset detection with a 50% criterion also reflects later portions of the LRP, and jackknifing may cause post-onset global noise. Yet, onset estimation based on the 50% criterion obviously was less affected by global noise than the regression-based method, as is indicated by the excellent performance of the 50%-jackknife method.

Perhaps, this apparent dissociation is related to a major disadvantage of regression-based onset detection, that is, treating all data points of the LRP rise equally. To achieve a more optimal fit to later portions of the LRP rise, a less optimal fit in the critical time range of true LRP onset may be occasionally accepted. This may result in relatively poor performance when variations in LRP shape come into play, which is the case with jackknifing. On the other hand, the major advantage of the 50%-relative criterion may be that it determines the LRP onset within a time range in which slope of the true LRP signal is maximal. In this area, large changes in amplitude are accompanied by only small changes in latency. Consequently, background noise added to the true LRP waveform affects LRP onset estimates less strongly than in the range of true LRP onset, close to 0 μV, where small changes in amplitude are accompanied by relatively large changes in latency.

True LRP onsets, and true correlation coefficients, were introduced by means of a very low absolute criterion (0.2 μV) which was used across all sets of simulations. Doubtless, a very low absolute criterion is most appropriate to determine *true* LRP onsets, given the definition of LRP onset, and that no preactivation needs to be assumed. It was only for their better reliability, not validity, that other scoring techniques like those using relative criteria were developed. For our purposes, fortunately, the 0.2-μV criterion could be applied because there was both no preactivation and very low noise in the true LRP waveforms.

However, one may suspect a bias favoring procedures that employ the same type of criterion, that is, an absolute, for LRP onset detection in the simulated, noisy LRPs. Given these considerations, it is all the more remarkable that the 50%-jackknife procedure performed most successfully, a criterion which is quite different from absolute criteria because it also depends on maximum LRP amplitude. This allows for two conclusions. First, it was not the correspondence to the criterion used to simulate a true correlation that made a certain criterion more effective in reproducing that correlation in the noisy data. Second, the reported consistent superiority of the 50%-jackknife method is independent of its sensitivity to LRP amplitude and/or slope. Consider, for example, two LRPs that start to deviate from zero at the same point in time, both reaching the same maximum amplitude, but at different latencies. According to the 50% criterion, the LRP reaching its maximum earlier will also be assigned an earlier onset. Note, however, that this feature of the 50% criterion, amplitude/slope sensitivity, cannot be responsible for the reported superiority of the 50%-jackknife method in the present correlative approach. This is because the true correlation was introduced using the 0.2-μV absolute criterion, which is completely independent of the maximum amplitude of a particular true LRP waveform drawn. Therefore, no systematic differences in LRP slope and/or amplitude did accompany true onsets, and no bias was created favoring scoring methods sensitive to amplitude/slope.

Although the present simulations did not suffer from amplitude confound, the situation in a single empirical study is somewhat different. Researchers who find a significant correlation between LRP onset according to a relative criterion and some external variable should be cautious regarding interpretations, because the correlation may have been introduced by amplitude rather than onset differences. With relative onset criteria, higher maximum LRP amplitude results in later LRP onset estimates. Thus, for reasons of safety, researchers should routinely try an absolute criterion together with the recommended 50%-relative criterion, to further corroborate their conclusion. In case of discrepant outcomes (50%-relative criterion produces a significant coefficient, x-μV absolute criterion does not), the (jackknifed) correlation between LRP amplitude and the external variable should be determined. If it is fairy low, this may justify preferring the relative criterion. Any substantial correlation between LRP amplitude and external variable, however, should be partialled out from the correlation involving LRP onset, before the latter is being interpreted.

Regarding the manipulation of interindividual variability in true LRP onset, based on the literature (Mordkoff & Gianaros, 2000) we expected to find a less pronounced superiority of jackknifing over single-subject procedures with increasing variability. For example, when a 50-ms effect in S-LRP onset had to be reproduced and variability was high, the single-subject 50% procedure more closely approached the true 50-ms difference and produced smaller standard deviation across simulation trials than the 50%-jackknife procedure (49.2 ms and 9.0 ms vs. 48.0 ms and 10.5 ms; see Table 4, high noise and high variability; Mordkoff & Gianaros, 2000). One may attribute this finding to the fact that higher interindividual LRP onset variability introduces stronger smearing, making jackknifed LRP onsets less reliable. However, in the present study, estimated correlation coefficients generally became *larger* with larger onset variability, and 50%-jackknife was again the most successful procedure. To be more specific, with a predefined r of 1.0, $n = 60$, and with medium noise, mean estimates obtained with the 50%-jackknife and the single-subject regression-based procedure were .18 and .10 (low variability), .36 and .19 (medium variability), and .46 and .26 (high variability), respectively (see Figures 3 and 4). Thus, the advantage of the 50%-jackknife procedure over the best single-subject procedure was about the same across variability conditions.

These results point to an important difference between the correlative approach and the approach of estimating mean differences in LRP onset. When jackknifed differences in LRP onset between two conditions are considered, the main effect of increasing variability within conditions is stronger smearing, making LRP onset estimates and estimates of onset differences *between* conditions less reliable. With correlations, on the other hand, the critical aspect is the correct reproduction of differences between individual LRP onsets *within* conditions. In other words, when LRP-onset variability between participants increases, differences between any two onsets tend to increase, which provides better contrast with background noise. Apparently, this effect of a greater signal-to-noise ratio predominates a potential detrimental effect related to stronger smearing because, in the present study, coefficients increased when between-subjects variability in LRP onset became higher.

Besides S-LRP latencies, research on interindividual differences in speed of information processing, specifically, motor processing, could also benefit from R-LRP latencies. For example, differences between extraverts and introverts regarding motor processing and motor activation have been assumed (Brebner & Cooper, 1978), and Rammsayer and Stahl (2004) found shorter R-LRP latencies in extraverts than introverts. There seems to be no reason why jackknifing could not also be successfully applied in correlative approaches involving R-LRP onsets. With regard to general shape of the waveform and intensity of noise added, one can assume S-LRP and R-LRP to be largely similar. If a combination of jackknifing and an appropriate relative criterion (50%) has proven to most reliably estimate the S-LRP onset, we can also expect that R-LRP onset can be successfully detected by a combination of jackknifing and an appropriate relative criterion (see Miller et al., 1998). Therefore, one may also reasonably assume that the investigation of correlations between R-LRP onsets and an external variable will benefit from the use of jackknife-based correlations.

So far, sources of unreliability of LRP onset detection and their impact on the estimated correlation coefficient have been discussed. However, in reality there are still other factors, making true correlations of 1.0 between LRP onset and an external variable rather unlikely. For example, reliability coefficients of personality scales rarely exceed .8, which sets an upper limit for realistic correlations between true LRP onset and personality. Further reduction will be caused by the fact that even if both measures were obtained with perfect reliability, common variability of LRP onsets and personality scores will always be smaller than 100%. Thus, a realistic estimate of the maximum true correlation may represent .7 or so. Given this true correlation, empirically found correlation coefficients will, on average, only range between .3 and .4 (with $n = 60$, smallest noise of 0.6 μV, and with the jackknifed 50%-relative criterion; see Figure 3). Table 3, however, tells us that statistical power is satisfactorily high (.88), that is, with 88% probability a significant correlation coefficient will be found that leads to the same conclusion as the true correlation. Because the above true correlation of .7 still represents a best-case scenario, researchers are also recommended to increase the likelihood of detecting true correlations smaller than .7 by trying to reduce LRP noise. Within participants, this can be achieved by obtaining a larger number of trials per hand. Fifty trials per hand resulted in LRP noise of 0.9 μV in the present study. Therefore, a number of at least 100 trials per hand should be run to approach our low-noise condition (0.6 μV). On the other hand, sample size should be sufficiently large, which is illustrated by the fact that, in the above example, statistical power was .69 with $n = 40$ as opposed to .88 with $n = 60$.

The advantage of using jackknifed data in correlative approaches may also be utilized outside personality research. Instead of personality scores, other dependent variables like RT, response force, or certain parameters of the evoked potential could serve as external variable. Correlating LRP onsets with RT or response force may provide additional insights into the relationship between different chronometric parameters. This opens the perspective that variables may be considered that cannot be experimentally manipulated as independent variables. For example, the correlation between jackknifed S-LRP onset and time markers in the evoked potential, like N100 or P300 latency, can be investigated using the present correlative approach, but not in factorial designs. Furthermore, even two jackknifed variables, for example, S-LRP onsets and R-LRP onsets, can now be correlated. Then, however, the sign of the coefficient must not be inverted.

A final cautionary note concerns a possible limitation of the present findings to experimental manipulations that do not provide pre-information. If pre-information regarding the correct response is available, one may expect true pre-onset deviations of the LRP from baseline. The present simulations employed data from a real experiment in which there was no pre-information. Consequently, pre-onset baseline was indeed close to zero for all true LRP waveforms. Our conclusions regarding the superiority of jackknifing in combination with a 50%-relative criterion may not apply to experiments that violate the assumption of zero pre-onset baseline.

To conclude, there seems to be a large number of areas within psychophysiological research where jackknife-based correlations may be successfully applied. Based on the present simulations, we can recommend the use of jackknifed S-LRP onsets determined with a 50%-relative criterion for correlative approaches, because this procedure performed consistently better than any other combination of estimation method and scoring technique. However, one of the methods suggested to control for LRP amplitude should be adopted.

REFERENCES

Brebner, J., & Cooper, C. (1978). Stimulus- or response-induced excitation: A comparison of the behavior of introverts and extraverts. *Journal of Research in Personality*, *12*, 306–311.

Coles, M. G. (1989). Modern mind-brain reading: Psychophysiology, physiology, and cognition. *Psychophysiology*, *26*, 251–269.

Coyle, S., Gordon, E., Howson, A., & Meares, R. (1991). The effects of age on auditory event-related potentials. *Experimental Aging Research*, *17*, 103–111.

Efron, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika*, *68*, 589–599.

Gratton, G., Coles, M. G., Sirevaag, E. J., Eriksen, C. W., & Donchin, E. (1988). Pre- and poststimulus activation of response channels: A psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 331–344.

Hackley, S. A., & Valle-Inclan, F. (1998). Automatic alerting does not speed late motoric processes in a reaction-time task. *Nature*, *391*, 786–788.

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371–375.

Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. [Cerebral potential changes in voluntary and passive movements in man: Readiness potential and reafferent potentials]. *Pflügers Archiv für die gesamte Physiologie*, *284*, 1–17.

Leuthold, H., Sommer, W., & Ulrich, R. (1996). Partial advance information and response preparation: Inferences from the lateralized readiness potential. *Journal of Experimental Psychology: General*, *125*, 307–323.

Meyer, D. E., Osman, A. M., Irwin, D. E., & Yantis, S. (1988). Modern mental chronometry. Special Issue: Event related potential investigations of cognition. *Biological Psychology*, *26*, 3–67.

Miller, J., & Hackley, S. A. (1992). Electrophysiological evidence for temporal overlap among contingent mental processes. *Journal of Experimental Psychology: General*, *121*, 195–209.

Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology*, *35*, 99–115.

Miller, R. G. (1974). The jackknife—A review. *Biometrika*, *61*, 1–15.

Mordkoff, J. T., & Gianaros, P. J. (2000). Detecting the onset of the lateralized readiness potential: A comparison of available methods and procedures. *Psychophysiology*, *37*, 347–360.

Osman, A. M., Bashore, T. R., Coles, M. G. H., Donchin, E., & Meyer, D. E. (1992). On the transmission of partial information: Inferences from movement-related brain potentials. *Journal of Experimental Psychology: Human Perception & Performance*, *18*, 217–232.

Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Posthuma, D., Mulder, E. J. C., Boomsma, D. I., & de Geus, E. J. C. (2002). Genetic analysis of IQ, processing speed and stimulus-response incongruency effects. *Biological Psychology*, *61*, 157–182.

Quenouille, M. H. (1949). Approximate tests of correlation in time-series. *Journal of the Royal Statistical Society*, *11*, 68–84.

Rammsayer, T. H., & Stahl, J. (2004). Extraversion-related differences in response organization: Evidence from lateralized readiness potentials. *Biological Psychology*, *66*, 35–49.

Sangals, J., Sommer, W., & Leuthold, H. (2002). Influences of presentation mode and time pressure on the utilisation of advance information in response preparation. *Acta Psychologica*, *109*, 1–24.

Schwarzenau, P., Falkenstein, M., Hoormann, J., & Hohnsbein, J. (1998). A new method for the estimation of the onset of the lateralized readiness potential (LRP). *Behavior Research Methods, Instruments, and Computers*, *30*, 110–117.

Smulders, F. T. Y., Kenemans, J. L., & Kok, A. (1996). Effects of task variables on measures of the mean onset latency of LRP depend on the scoring method. *Psychophysiology*, *33*, 194–205.

Stelmack, R. M., Houlihan, M., & McGarry-Roberts, P. A. (1993). Personality, reaction time, and event-related potentials. *Journal of Personality & Social Psychology*, *65*, 399–409.

Ulrich, R., & Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology*, *38*, 816–827.

Ulrich, R., Leuthold, H., & Sommer, W. (1998). Motor programming of response force and movement direction. *Psychophysiology*, *35*, 721–728.

Ulrich, R., Rinkenauer, G., & Miller, J. O. (1998). Effects of stimulus duration and intensity on simple reaction time and response force. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 915–928.

Vaughan, H. G., Costa, L. D., & Ritter, W. (1968). Topography of the human motor potential. *Electroencephalography & Clinical Neurophysiology*, *25*, 1–10.

Walhovd, K. B., & Fjell, A. M. (2002). One-year test-retest reliability of auditory ERPs in young and old adults. *International Journal of Psychophysiology*, *46*, 29–40.

(Received August 8, 2003; Accepted June 18, 2004)

APPENDIX

Proof A

The purpose of this proof is to provide Equation 1, that is, that the correlation coefficient r_{jx} between the scores x_i of n participants ($i = 1 \ldots n$) on any nonjackknifed variable x and scores j_i on a second, jackknifed variable equal -1 times the correlation coefficient r_{ox} between the n original, single-subject scores o_i and scores x_i. Because it is required for this proof, in Proof B we show that the standard deviation of jackknifed scores equals the standard deviation of the respective single-subject scores, divided by $(n - 1)$.

$$r_{jx} = -r_{ox}. \tag{1}$$

The correlation coefficients r_{jx} and r_{ox} can be denoted as the expectation of the product of z values of o and x and j and x, respectively. Thus, we can write

$$E(z_j z_x) = -E(z_o z_x). \tag{2}$$

Equation 2 holds if we can provide that for each subject i, the z value of a single-subject score o_i equals the inverse of the z value of the jackknifed score j_i (including all $n - 1$ subjects but i from the given sample, n), which reads

$$z_{ji} = -z_{oi}. \tag{3}$$

Let $\bar{J}$ (s_j) and $\bar{O}$ (s_o) be the grand averages (standard deviations) of jackknifed scores j_i and single-subject scores o_i, respectively, obtained from n participants. Then, Equation 3 results in

$$\frac{j_i - \bar{J}}{s_j} = -\frac{o_i - \bar{O}}{s_o}. \tag{4}$$

After multiplying, one yields

$$(j_i - \bar{J}) \cdot s_o = -(o_i - \bar{O}) \cdot s_j. \tag{5}$$

The jackknifed score j_i of any subject i is defined by the following equation:

$$j_i = \frac{\sum_{v=1}^{n} o_v}{n - 1}, \text{ with } v \neq i, \tag{6}$$

where $v = 1 \ldots n$. The grand average $\bar{J}$ of all n jackknifed scores j_i is defined as

$$\bar{J} = \frac{\sum_{i=1}^{n} j_i}{n}. \tag{7}$$

Inserting Equation 6 results in

$$\bar{J} = \frac{\sum_{i=1}^{n} \left(\frac{\sum_{v=1}^{n} o_v}{n-1} \right)}{n} = \frac{\sum_{i=1}^{n} \left(\sum_{v=1}^{n} o_v \right)}{n(n-1)}, \text{ with } v \neq i. \tag{8}$$

In the numerator of Equation 8, each o_i is added to the sum $(n - 1)$ times. Hence,

$$\bar{J} = \frac{(n-1)\sum_{i=1}^{n} o_i}{n(n-1)} = \frac{\sum_{i=1}^{n} o_i}{n} = \bar{O}. \tag{9}$$

Because the means of j_i and o_i are identical (Equation 9), we transform Equation 5 into

$$j_i \cdot s_o + o_i \cdot s_j = \bar{O} \cdot (s_j + s_o). \tag{10}$$

For the standard deviation of jackknifed scores s_j and the standard deviation of single-subject scores s_o, the following simple relation is true (see Proof B):

$$s_o = s_j(n - 1). \tag{11}$$

Inserting Equation 11 into Equation 10 and dividing by s_j results in

$$(n - 1)j_i + o_i = \bar{O} \cdot n. \tag{12}$$

If Equation 6 is inserted into Equation 12, one yields

$$(n - 1)\frac{\sum_{v=1}^{n} o_v}{n - 1} + o_i = \bar{O} \cdot n, \text{ with } v \neq i, \tag{13}$$

which results in

$$\frac{\sum_{i=1}^{n} o_i}{n} = \bar{O}. \tag{14}$$

Equation 14 is always true, which completes the proof.

Proof B

The purpose of the second proof is to provide the assumed relation (Equation 11) between s_j and s_o, the standard deviations of jackknifed and original scores, respectively. The squared standard deviation, that is, the variance, of the jackknifed scores s_j^2 is defined as

$$s_j^2 = \frac{\sum_{i=1}^{n} (j_i - \bar{J})^2}{n - 1}. \tag{15}$$

After multiplying Equation 6 by $(n - 1)$ and adding o_i to both sides, we can write

$$\sum_{v=1}^{n} o_v + o_i = j_i \cdot (n - 1) + o_i, \text{ with } v \neq i, \tag{16}$$

From Equation 9 we can derive

$$\sum_{i=1}^{n} o_i = \bar{O} \cdot n. \tag{17}$$

Because the left-hand side of Equation 16 equals the left-hand side of Equation 17, one obtains

$$j_i(n - 1) + o_i = \bar{O} \cdot n, \tag{18}$$

and

$$j_i = \frac{\bar{O} \cdot n - o_i}{n - 1}. \tag{19}$$

Subtracting $\bar{J}$ on both sides results in

$$j_i - \bar{J} = \frac{\bar{O} \cdot n - o_i}{n-1} - \bar{J}. \tag{20}$$

Given that the grand average of jackknifed values $\bar{J}$ equals the grand average $\bar{O}$ of single-subject scores (Equation 9), it follows that

$$j_i - \bar{J} = \frac{\bar{O} \cdot n - (n-1) \cdot \bar{O} - o_i}{n-1}, \tag{21}$$

and

$$j_i - \bar{J} = \frac{\bar{O} - o_i}{n-1}. \tag{22}$$

Inserting Equation 22 into the numerator of Equation 15, one obtains

$$s_j^2 = \frac{\sum_{i=1}^{n} (\bar{O} - o_i)^2}{(n-1)^2 \cdot (n-1)}. \tag{23}$$

Because of the definition of the variance of o_i, Equation 23 results in

$$s_j^2 = \frac{s_o^2}{(n-)^2}. \tag{24}$$

It follows immediately that

$$s_j = \frac{s_o}{n-1}. \tag{25}$$

and this completes the proof.