

INDIVIDUAL ANALYSIS OF LATERALITY DATA

MARC BRYSSBAERT and GÉRY D'YDEWALLE

Department of Psychology, University of Leuven, B-3000 Leuven, Belgium

(Received 27 April 1989; accepted 16 January 1990)

Abstract—Graphical and statistical analyses are presented that allow one to check for an individual subject whether the performance during a session is stable, whether the difference between the left and the right visual half-field is significant, and whether the performance is uniform over different sessions. Analyses are given for accuracy data and for latency data. Though the analyses are described for a visual half-field experiment, they can easily be adapted for other laterality tasks.

INDIVIDUAL ANALYSIS OF LATERALITY DATA

DESPITE 30 years of intensive laterality research it still is rather difficult to set up a visual half-field task for determining cerebral dominance. One reason is the lack of reliability data, the other is the difficulty in finding information about individual data analysis. The reliability problem will only be mentioned in passing, as it has been dealt with elsewhere [4]. Here, we will mainly be concerned with the question of finding standards for individual assessment.

A distinction must be made between accuracy and latency data, though both variables may be assessed in the same experiment. For the individual analysis of accuracy data, we will mainly work with the lambda index (discussed below), because prior experimentation [4, 6] has shown that it correlates highly with other possible indices, and its theoretical foundation is well documented by SPROTT and BRYDEN [2, 18]. For the analysis of latency data, the point-biserial correlation [12] will be used because of its higher reliability than the mere difference index [4] and because of its more elegant statistical properties.

Two analyses will be described that can be used for both accuracy and latency data, and that complement each other very well. The first one is a graphical analysis, the second one, a statistical analysis. A graphical analysis is given because a lot of information that is difficult to grasp with statistics, can easily be represented in a figure. Some of the statistical methods have already been introduced in the neuropsychological literature before [2, 12, 18], but will be repeated here in order to give a full picture. The analyses are described for visual half-field (VHF) tasks, but can easily be extended to other laterality experiments.

ACCURACY DATA

Graphical analysis

If the stimuli are well randomized and if the subject performs properly, then the number of stimuli recognized in one VHF should be more or less linearly related to the number of stimuli presented in that VHF. Thus, the pattern of results in one VHF can ideally be represented by a straight line with a slope equal to the mean proportion of the correctly

identified stimuli. Figure 1(A) shows the ideal situation for a left hemisphere dominance situation; Fig. 1(B) gives the actual performance of a good subject (left hemisphere dominance, words as stimuli).

Departures from the straight line are caused by fluctuations in attention on the subject's side or by grouping of easier and more difficult stimuli. They are not influenced by the random assignment of stimuli to the left and right VHF.

Three situations, apart from the good one, deserve special attention. First, a departure of both cumulative functions (VHFs) in the same direction very probably points to increases or decreases of the subject's overall attention during the experimental session. Figure 2(A) gives an example of a subject who showed a decreased attention at the beginning of the session. Second, a departure of the cumulative function in opposite directions indicates that the subject directs his attention to one VHF. In Fig. 2(B), this is successively to the left and to the right VHF. Finally, a departure of one cumulative function without an effect on the other is most likely to be caused by a clustering of easy or difficult stimuli on one side of the fixation location.

Unfortunately, neither graphical nor statistical analysis will detect more subtle strategies adopted by the subject (e.g. always paying attention to the VHF in which the previous stimulus appeared), unless they are especially looked for by the experimenter (e.g. by comparing the proportion of correct stimuli presented to the same VHF as the previous stimulus, with the proportion of correct stimuli presented to the opposite VHF).

Statistical analysis

Once we have noted that the subject displays a stable pattern of performance, two other questions become important. Is there a reliable difference between the left visual field (LVF) and the right visual field (RVF), and is the difference uniform over successive experimental sessions?

The lambda index proposed by SPOTT and BRYDEN [2, 18] allows a rather simple answer to these questions. It has been shown that the (natural) logarithm of the number of correct responses divided by the number of incorrect responses is approximately normally distributed, the variance being equal to the sum of the reciprocals of the number of correct and incorrect responses. Thus,

$$\log(n_+/n_-) \sim .N(\log(n_+/n_-), 1/n_+ + 1/n_-) \quad (1)$$

n_+ = the number of correct responses,

n_- = the number of incorrect responses.

Equation (1) forms the basis to check whether the difference between the number of correct responses in the RVF differs significantly from the number of correct responses in the LVF. Because the index of Equation (1) is normally distributed for the RVF and the LVF, the difference between the indices of both VHFs (further called lambda index) will be normally distributed too, the variance being equal to the sum of the variances of the separate indices. Thus,

$$\begin{aligned} \text{lambda} &= \log(n_{+r}/n_{-r}) - \log(n_{+l}/n_{-l}) \\ &\sim .N(\text{lambda}, 1/n_{+r} + 1/n_{-r} + 1/n_{+l} + 1/n_{-l}). \end{aligned} \quad (2)$$

To see whether a lambda is significantly different from zero, a simple look at a standard

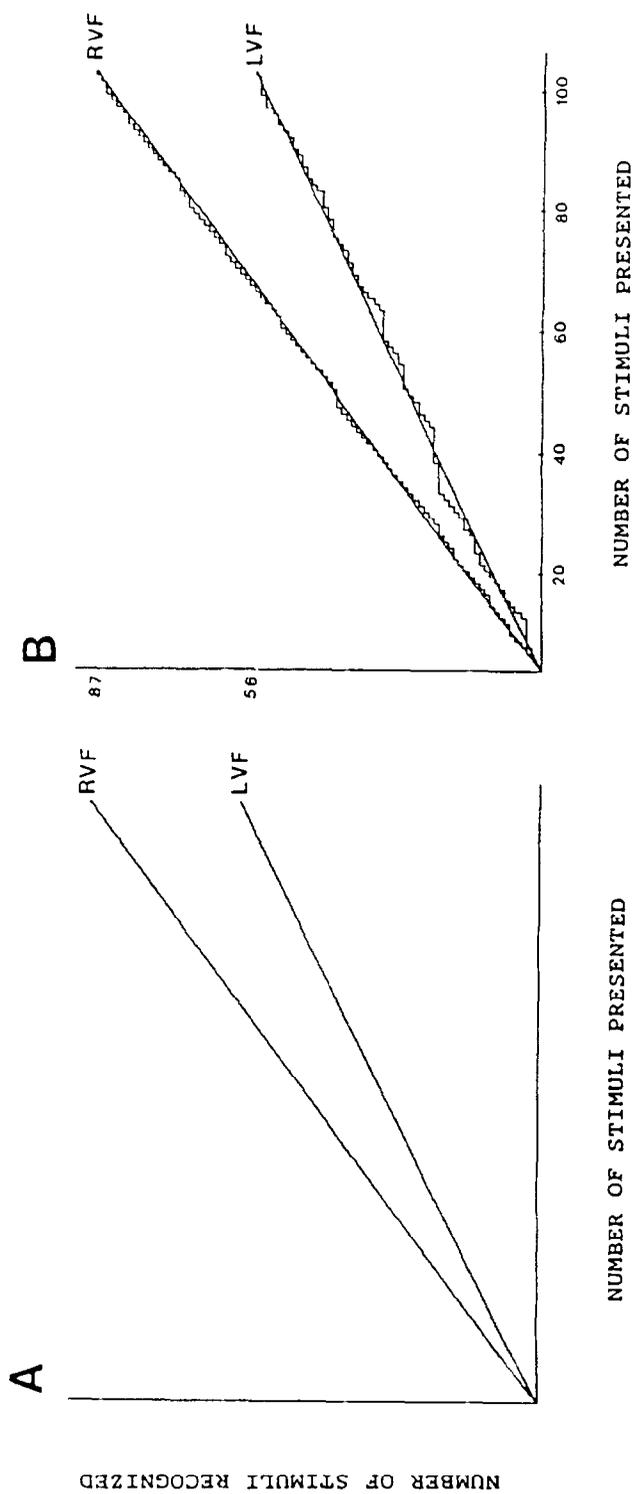


Fig. 1. Graphical analysis of accuracy data: (A) ideal situation for left hemisphere dominance, (B) performance of a good subject for 100 five-letter words presented to the LVF and RVF.

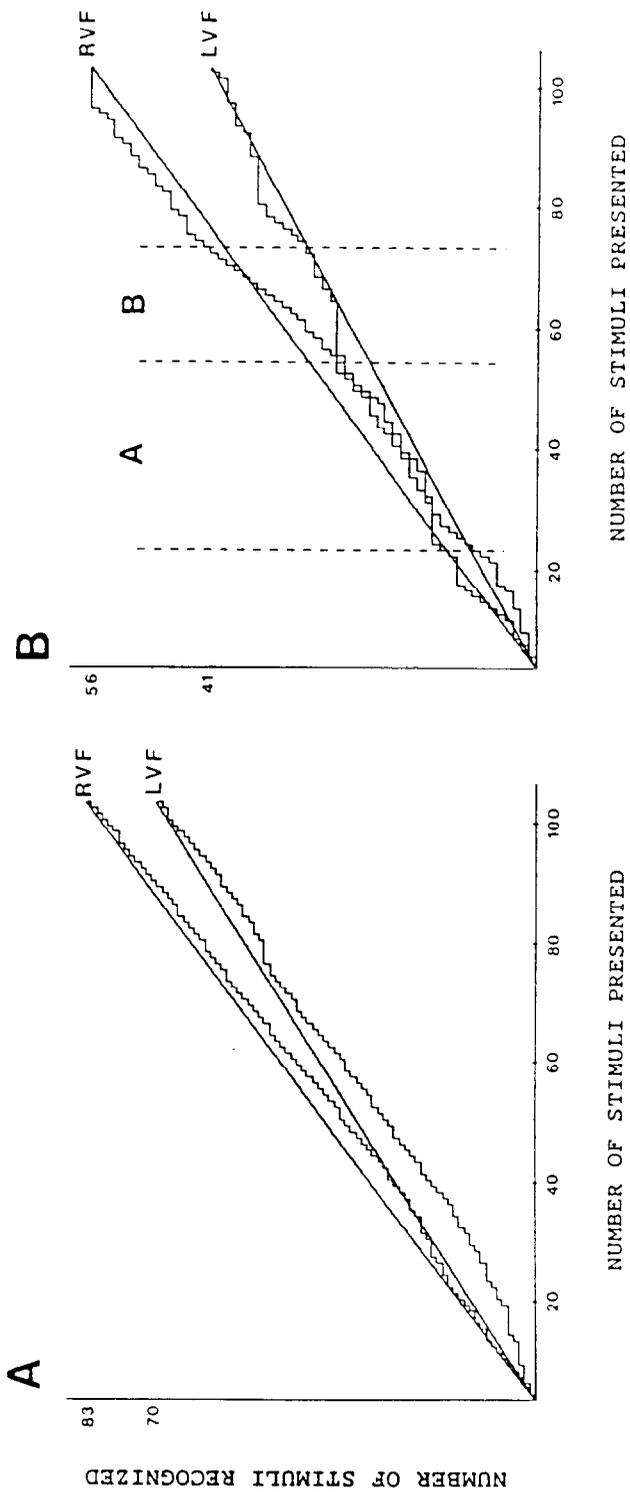


Fig. 2. Graphical analysis of accuracy data: (A) subject showing decreased attention at the beginning of the session, (B) subject showing a shift of attention, first to the LVF (A), then to the RVF (B).

normal table suffices. Similarly, confidence intervals can be computed, as has been done in Fig. 3 for a subject who finished five series of approximately 50 four-letter words in each VHF, and five series of approximately 50 five-letter words in each VHF (for more details, see [4]).

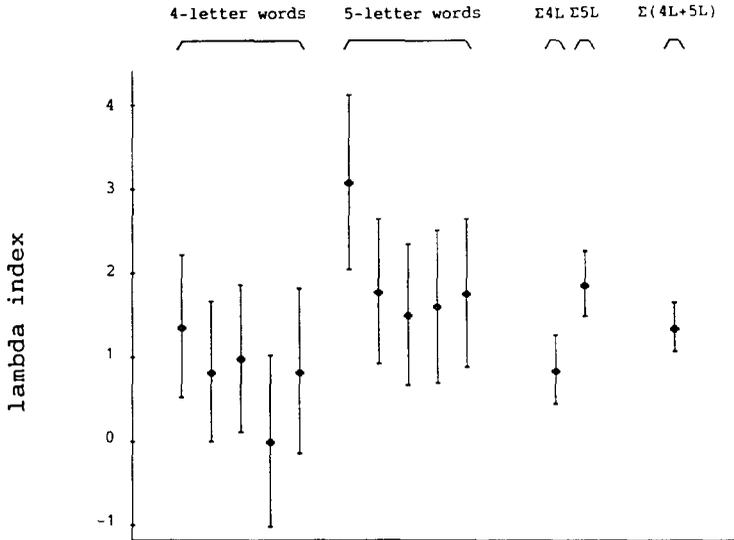


Fig. 3. Lambda indices and 95% confidence intervals for a VHF study in which five series of four-letter words and five series of five-letter words were administered (data from one subject). Data for the separate series, as well as for the sum of the four-letter word series ($\Sigma 4L$), the sum of the five-letter word series ($\Sigma 5L$), and the sum of all series ($\Sigma(4L + 5L)$).

A remarkable aspect of Fig. 3 is the size of the confidence intervals around the lambdas based on a single series, a problem we did not appreciate that well before we drew the figure. It demonstrates how cautiously individual results must be interpreted if they are not based on sufficient data (i.e. compare the confidence intervals of the single series with the confidence interval of their sum, Fig. 3).

Because of the statistical foundation, it is possible to estimate how large the difference between two lambdas must be to reach significance. The smallest variance a lambda can have is achieved in a situation in which half of the stimuli in each VHF are recognized. For instance, if we present 50 stimuli in each VHF, the smallest variance is $1/25 + 1/25 + 1/25 + 1/25 = 0.16$. This means that the difference between two lambdas must exceed 1.11 to be significant at the 0.05 level. For,

$$\begin{aligned}
 &(\lambda_1 - \lambda_2) / \sqrt{\text{var}_{L_1} + \text{var}_{L_2}} > 1.96 \\
 &(\lambda_1 - \lambda_2) / \sqrt{0.16 + 0.16} > 1.96 \\
 &(\lambda_1 - \lambda_2) > 1.96 \times 0.57.
 \end{aligned}
 \tag{3}$$

The difference will always have to be larger than the minimal value, because it is impossible to have at the same time a difference between two lambdas and a minimal variance for both lambdas. The size of the confidence region surely has to be taken into account, if correlations over subjects between lambdas and some other variable are investigated (see e.g. [6]). Table 1 gives the minimal difference for some numbers of stimuli administered to each VHF

to reach the 5% significance level. Personally, we would feel very uncertain if lambda indices were based on less than 50 stimuli per VHF, also because the accuracy of the normal approximation of the lambda index depends on the number of stimuli administered.

The second question we asked at the beginning of the section, was how we can appreciate the homogeneity of lambdas obtained over different sessions. For instance, we might want to know whether the 10 lambdas of Figure 3 are homogeneous. SPOTT and BRYDEN [2, 18] describe the following procedure: All lambdas can be translated into approximately standard normal quantities (u_i). Therefore, assuming that all r replications are independent experiments, the u_i^2 are independent chi-square variates with one degree of freedom, so that Σu_i^2 is a chi-square variate with r degrees of freedom. Under H : $\lambda_{i1} = \lambda_{i2} = \dots = \lambda_{ir} = \lambda$, the quantity Σu_i^2 is a function of the common lambda, an estimate of which is:

$$L_{est} = (\Sigma L_i / s_i^2) / (\Sigma 1 / s_i^2), \quad (4)$$

L_{est} = estimate of the grand lambda,

L_i = lambda of the i th replication,

s_i^2 = variance of lambda of the i th replication.

From the foregoing, it follows that a good statistic to test the homogeneity of lambdas is:

$$\Sigma u_i^2 = \Sigma (L_i - L_{est})^2 / s_i^2, \quad (5)$$

which has an approximate chi-square distribution with $r - 1$ degrees of freedom, one degree of freedom being lost in estimating L_{est} . Table 2 shows the results of such an analysis for an experiment [4] in which approx. 50 words were presented to the LVF and RVF. Fourteen subjects participated, half of whom were right-handed (s1-s7), and half left-handed (s8-s14). Each subject got five replications of a series of four-letter words (series 1-5) and five replications of a series of five-letter words (series 6-10). The data of the first subject were used to create Figure 3.

If the individual L_i are significantly different, some multiple test procedure can be used to find out which lambdas are different from each other. MARASCUILO [14] proposes a chi-square test, which allows a *post hoc* test for every possible comparison. A comparison is significant if:

$$\text{abs}(\Sigma w_i L_i) / \text{sqrt}(\Sigma w_i^2 s_i^2) > \text{sqrt}(\text{chi-sq}_{r-1}(1-\alpha)), \quad (6)$$

in which the w_i are the contrast coefficients such that $\Sigma w_i = 0$, and r is the number of lambdas involved in the study. For instance, YOUNG and ELLIS' [19] claim that five-letter words yield larger laterality indices than four-letter words is confirmed for subject 1 in Table 2 (also see Fig. 3) if:

$$\text{abs}(L_{5L} - L_{4L}) / \text{sqrt}(s_{5L}^2 + s_{4L}^2) > \text{sqrt}(\text{chi-sq}_9(0.95)), \quad (7)$$

in which $L_{5L} = 1/5 \times (L_6 + L_7 + L_8 + L_9 + L_{10})$

$$= (3.1 + 1.8 + 1.5 + 1.6 + 1.8) / 5$$

$$= 1.96$$

$$L_{4L} = 1/5 \times (L_1 + L_2 + L_3 + L_4 + L_5)$$

$$= (1.4 + 0.8 + 1.0 + 0.0 + 0.8) / 5$$

$$= 0.8$$

$$s_{5L}^2 = 1/25 \times (s_6^2 + s_7^2 + s_8^2 + s_9^2 + s_{10}^2)$$

$$= (0.28 + 0.19 + 0.18 + 0.22 + 0.20)/25$$

$$= 0.04$$

$$s_{4L}^2 = 1/25 \times (s_1^2 + s_2^2 + s_3^2 + s_4^2 + s_5^2)$$

$$= (0.19 + 0.18 + 0.20 + 0.27 + 0.25)/25$$

$$= 0.04$$

chi-sq₉(0.95) = 16.9.

Table 1. Minimum difference between two lambda indices needed to be reliable at 5% for different numbers *N* of stimuli presented to each of the two VHF's.

<i>N</i>	Minimum difference
5	3.51
25	1.57
50	1.11
250	0.50
500	0.35
2500	0.16
5000	0.11

Table 2. Lambda indices of five replications of a series of four-letter words (Series 1-5), and of a series of five-letter words (Series 6-10). Grand lambda calculated with the use of Equation (4), probability of grand lambda being equal to zero calculated with Equation (4) and (8), probability of the 10 lambdas being homogeneous calculated with Equations (4) and (5). Underlined lambdas differ at a 5% level from each other [Equations (4) and Table 3].

	1	2	3	Series		6	7	8	9	10	Grand Lambda	<i>P</i> (=0)	<i>P</i> (homog)
				4	5								
s1	1.4	0.8	1.0	<u>0.0</u>	<u>0.8</u>	<u>3.1</u>	1.8	1.5	1.6	1.8	1.37	0.000	0.007
s2	1.1	1.1	1.0	1.6	1.2	2.2	1.5	2.5	2.9	1.5	1.57	0.000	0.310
s3	3.1	2.2	1.6	2.2	2.4	3.0	2.8	3.0	2.6	2.8	2.53	0.000	0.555
s4	0.0	0.8	2.4	0.4	1.2	1.9	1.2	1.1	1.1	1.1	1.07	0.000	0.276
s5	0.8	1.3	1.6	1.8	1.6	0.5	3.1	2.1	1.8	1.6	1.49	0.000	0.033
s6	0.2	1.2	1.1	0.5	1.9	0.5	1.7	1.8	0.6	0.6	0.95	0.000	0.028
s7	0.8	1.2	0.9	1.5	1.6	2.1	2.4	2.5	2.5	1.6	1.67	0.000	0.038
s8	1.5	2.2	1.4	1.6	2.1	1.4	2.7	3.1	3.6	1.9	2.00	0.000	0.030
s9	0.3	1.6	1.1	0.2	0.6	-0.4	0.2	0.6	0.6	0.1	0.54	0.000	0.170
s10	1.1	1.2	0.7	1.5	0.8	2.2	3.1	2.4	2.0	3.1	1.65	0.000	0.000
s11	0.1	0.4	0.2	-0.7	0.3	-0.2	-0.1	0.5	0.2	0.3	0.11	0.424	0.850
s12	-0.2	1.4	0.2	1.0	0.8	1.1	1.6	0.7	1.0	0.8	0.83	0.000	0.168
s13	1.2	0.0	-0.4	1.1	0.8	2.2	1.3	1.6	1.8	1.4	1.11	0.000	0.003
s14	1.4	2.0	0.8	0.8	1.7	2.0	2.1	2.1	2.0	1.6	1.62	0.000	0.281

The difference between five- and four-letter words divided by their joint SD (i.e. 4.0) lies quite close to the critical value (i.e. 4.1), so that YOUNG and ELLIS' [19] statement receives some support from the data of subject 1 in Table 2.

However, because Marascuilo's procedure determines the critical value in such a way that the joint probability of a Type-1 error is smaller than or equal to α for the total set of all possible contrasts, it will be conservative if we are interested in a limited set of comparisons, for instance if we only are interested in the set of pairwise comparisons (Marascuilo's procedure yields but one significant pairwise comparison in Table 2, that between the two most extreme lambdas of subject 1). Therefore, if we want to limit the *post hoc* comparisons to pairwise contrasts, it is more interesting to use a stepwise method, such as HOLM's [9, see also 8] sequentially rejective procedure. In this procedure, the *P*-value of all $r \times (r - 1)/2$ pairwise contrasts are tabulated in an ascending order and compared with the critical value $\alpha(k - i + 1)$, in which *k* is the total number of pairwise contrasts and *i* is the rank number of the contrast to be compared. The procedure rejects null hypotheses of no contrast as long as the *P*-values of the contrasts are below the critical values. Whenever the *P*-value exceeds the critical value, the procedure must be stopped, even if on later occasions the *P*-value again falls below the critical value. Table 3 shows how the procedure works for the data of s1 in Table 2 (see also Fig. 3). HOLM [9] proved that the multiple alpha-level of his procedure does not exceed α , a property which the more familiar multiple range test procedures of the Newman-Keuls type do not maintain. Table 3 contains the results of Holm's procedure for pairwise comparisons.

Table 3. Review table of the sequentially rejective test for subject 1 of Table 2.

<i>i</i>	L_a	L_b	<i>u</i>	<i>P</i> (<i>u</i>)	$\alpha/(k - i + 1)$
1	3.08	0.00	4.15	0.00005	< 0.00111
2	3.08	0.82	3.33	0.00086	< 0.00114
3	3.08	0.83	3.09	0.00200	> 0.00116
					→ stop testing
4	3.08	0.98	3.03	0.00244	0.00119
5	1.79	0.00	2.64	0.00831	0.00122
⋮	⋮	⋮	⋮	⋮	⋮
45	0.83	0.82	0.02	0.98783	0.05000

i = rank number of pairwise comparison, ranged in ascending order with respect to the *p*-value.
 L_a, L_b = lambdas involved in the pairwise comparison.
u = standard normal value of the difference between L_a and L_b (see Equation (2)).
p(*u*) = two-tailed significance probability associated with *u*.
 α = multiple alpha-level.
k = total number of pairwise comparisons involved in the analysis (here: $(10 \times 9)/2 = 45$).

A significant difference from zero of the overall lambda (L_{est}) can be tested with the use of the standard normal distribution, for L_{est} is normally distributed with variance

$$s_{L_{est}}^2 = 1/\sum(1/s_i^2). \tag{8}$$

independent [18, pp. 450–451]. That is, being correct on one trial must not be affected by the correctness of the preceding trials. This assumption will better be met if stimulus presentation is thoroughly randomized and if adequate fixation of the subject to the middle of the visual field is ensured. A careful researcher, however, might in addition want to check serial independence *post hoc* by calculating autocorrelations in the obtained data [10, pp. 287–290]. The graphical analysis presented above will be of some help too, in order to see whether there are dependencies in the correctness of the subject's responses.

LATENCY DATA

Graphical analysis

A similar type of graph as for the accuracy data (Fig. 1) can be drawn to check whether the subject's performance fluctuated during the session. A change, however, must be made because in most experiments only reaction times for correct answers are interesting, and the number of good responses differs between VHF's. Therefore, the scale of the abscissa for the LVF and RVF will be different, and the reaction latencies must be divided by the number of correct responses.* This leads to curves ranging on the ordinate from zero (at the beginning of the experiment) to the average response latency per VHF (when the reaction time for the last correct response has been entered). Figure 4 shows such curves for a good subject who finished an experiment in which 100 words were presented to the LVF and RVF.

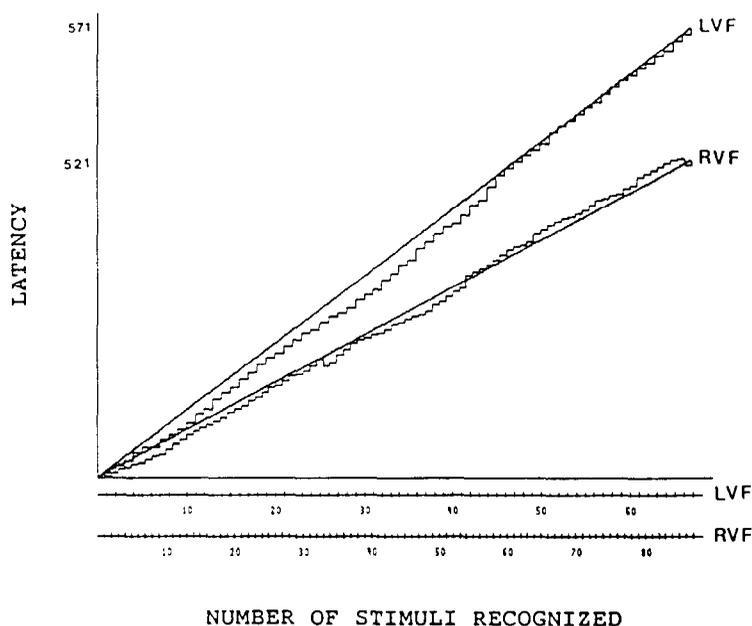


Fig. 4. First graphical analysis of latency data. Performance of a good subject for 100 five-letter words presented to LVF and RVF.

*To get a clearer distinction between the curves of the LVF and RVF, it is better not to take the cumulative reaction latencies, but the cumulative reaction latencies diminished by a constant, thus $\Sigma(\text{latency} - x)$, where $x \leq$ minimal reaction latency.

Another interesting graph is obtained by plotting the latency on the abscissa and the cumulative proportion of latencies on the ordinate. The rationale for this approach is that every difference between the mean reaction time of RVF and LVF must be characterized by a shift of the complete distribution to be genuine. Differences due to an excess of very low or very high reaction latencies are always more or less questionable. A good performance should be marked by two well separated S-shaped curves. The curves will slightly differ in form because reaction time distributions are known to be positively skewed and to have a positive correlation between mean and variance [13].

An example of how the graph should look for stimuli preferentially processed in the left hemisphere, is given in Fig. 5 (Fig. 5(A) gives the ideal situation, Figure 5(B) the actual performance of a good subject for 100 words presented to the RVF and LVF).

An advantage of the graphical display is that it is rather insensitive to outliers at the lower or the higher end of the scale. It also gives a more complete picture than mean or median latencies which, because of the specific distribution of reaction times, sometimes are biased estimators of the population values [16].

Statistical analysis

LEVY [12] holds that the variability in reaction latencies is inherent to hemispheric specialization and therefore proposes to use the point-biserial correlation coefficient (r_{pb}) as a better laterality index than the mere difference between the mean latency for the LVF and the RVF. An empirical VHF study in which two series of 100 words were presented five times [4] added support to Levy's position in that test-retest reliability was higher for the r_{pb} index than for the difference index.* It also pointed to the high correlation between the r_{pb} and the difference index. Therefore, for the rest of the article we will confine ourselves to the point-biserial correlation coefficient.

The r_{pb} is calculated with the use of Equation 9.

$$r_{pb} = \frac{RT_L - RT_R}{s} \times (P_L P_R)^{1/2}, \quad (9)$$

RT_L = mean reaction time for LVF,

RT_R = mean reaction time for RVF,

P_L = proportion of stimuli presented in LVF,

P_R = proportion of stimuli presented in RVF,

s = SD of all reaction times.

Though the exact distribution of r_{pb} can be determined, DAS GUPTA [5] gives a variance-stabilizing transformation of r_{pb} , that approximately leads to a normal distribution

*The relation between the reliability of the difference score (or the r_{pb} coefficient) and the reliability of the measures in the VHFs depends on the covariation between the measures of the VHFs ([12], pp. 467-469). If the covariation is positive, the reliability of the difference (r_{pb}) is lower than the reliability of the measures; if the covariation is negative, the reliability of the difference (r_{pb}) is higher; and if the covariation is zero, reliabilities of laterality indices and raw measurements are equal.

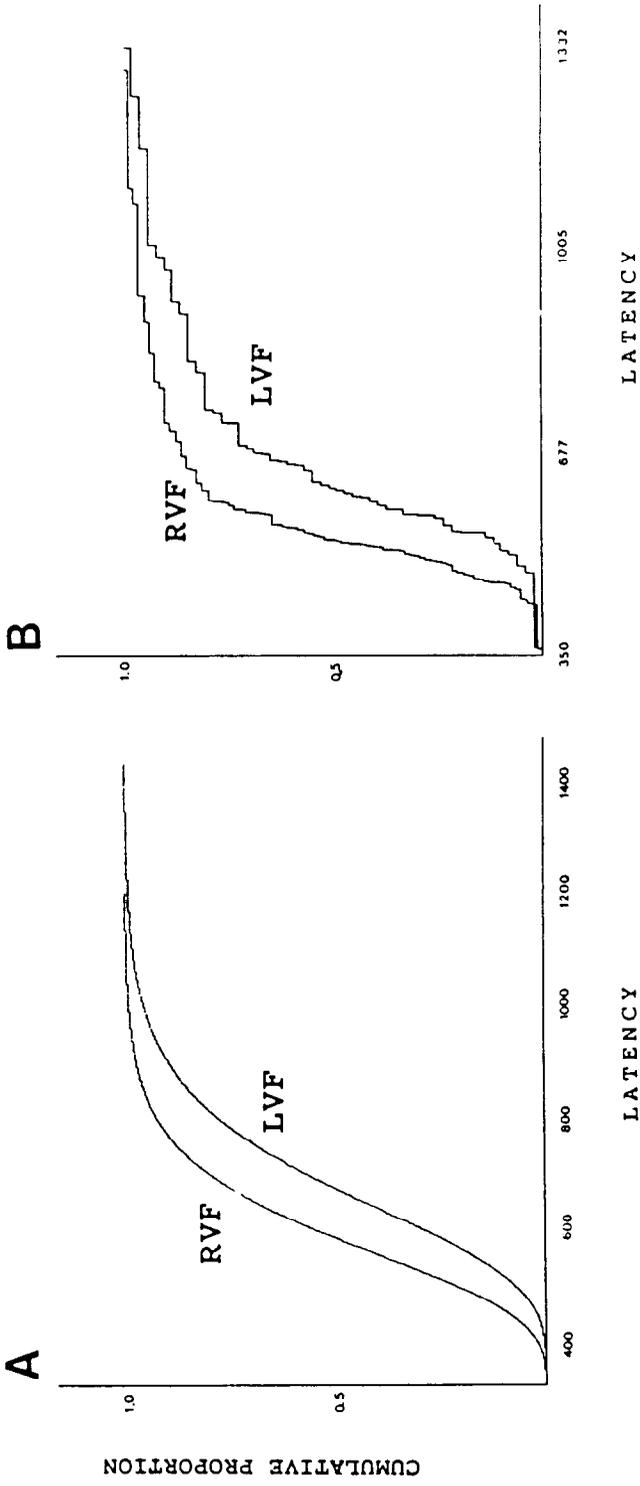


Fig. 5. Cumulative proportion of naming latencies for 100 five-letter words presented to LVF and RVF: (A) ideal situation simulated with two lognormal distributions, (B) performance of a good subject.

and therefore allows the same analyses as described for the lambda index. The transformation is given by:

$$f(r_{pb}) = \frac{1}{(A+B)^{1/2}} \times 0.5 \times \log[(1+C)/(1-C)], \quad (10)$$

$$A = (P_R + P_L K)/(P_L + P_R K),$$

$$K = s_R^2/s_L^2,$$

s_R^2 = variance of RVF,

s_L^2 = variance of LVF,

$$B = \frac{1}{4P_L P_R} - 1.5 \times \frac{P_R + P_L K}{P_L + P_R K} + 0.75 \times \frac{(1-K)^2}{(P_L + P_R K)^2} \times P_L P_R,$$

$$C = \frac{r_{pb} \times (A+B)^{1/2}}{(A+B \times r_{pb}^2)^{1/2}}.$$

The variance of $f(r_{pb})$ is the reciprocal of the total number of stimuli administered ($1/N$).

A difficulty with Das Gupta's transformation is the explicit assumption it makes that the latencies are normally distributed in the RVF and LVF. As is well known [13] and indicated above, response times are not characterized by a normal but a positively skewed distribution. We have done extensive simulations to investigate the effect of the departure using a lognormal distribution as an approximation, and the general result is that a positively skewed distribution leads to a small reduction of the power of the test. The extent of the reduction depends on the size of r_{pb} (the more it deviates from zero, the greater the reduction), and the ratio of the number of correct responses in the RVF and the LVF (the reduction is smallest when the ratio is close to one), but not (much) on the total number of correct responses provided $n > 50$.

Three attitudes might be adopted concerning the loss of power. First, a better transformation might be sought; second, the logarithm of the response latencies can be used as a dependent variable; and third, the loss may be accepted as such. In this manuscript we adopt the last attitude. Finding a better transformation would make things rather complicated [5] and might be well beyond the wit of most laterality researchers (including the authors'). Taking the logarithm of the response times implies a double transformation to be made, so that the final index loses much of its understanding, certainly if we take the benefits into consideration. The conservatism will be larger for substantial r_{pb} 's and/or differences in accuracy between the two VHF's, both situations in which statistics are not needed that much to make firm conclusions about the laterality pattern. Table 4 contains the results of some simulations and gives an idea of the extent of information lost in typical situations. As can be seen, the drop in power for these less extreme situations is well within the limits.

Testing the significance of an r_{pb} is carried out analogously to the testing of a lambda index. Because $f(r_{pb})$ is normally distributed with variance $1/N$, a standard normal table can be used to test for significance and to calculate 95% confidence intervals. Figure 6 shows the results of the latter analysis for the correct responses of the subject mentioned in Fig. 3.

Again, the size of the confidence regions is noteworthy and urges to present sufficient stimuli to all subjects. Equation (11) gives r_{pb} as a function of $f(r_{pb})$, needed to determine confidence intervals.

Table 4. Comparison of normal and lognormal distributions with respect to the normality assumption of Das Gupta's transformation of r_{pb} . μ = the expected value of the distribution, σ = SD. N = the number of observations taken from the distribution, r_{pb} = the point-biserial correlation that is to be expected, and $P(u)$ = one-tailed significance probability associated with standardized value u . Data based on 20,000 r_{pb} 's per distribution.

Distributions	μ	σ	N	r_{pb}	$P(-2.58)$	$P(-1.96)$	$P(1.96)$	$P(2.58)$
$715.4 + u \times 158.2$	715.4	158.2	24	0.31	0.005	0.024	0.034	0.009
$622.0 + u \times 129.5$	622.0	129.5	38					
$200 + \exp(6.2 + u \times 0.3)$	715.4	158.2	24	0.31	0.004	0.020	0.023	0.005
$200 + \exp(6.0 + u \times 0.3)$	622.0	129.5	38					
$715.4 + u \times 158.2$	715.4	158.2	48	0.31	0.005	0.025	0.032	0.008
$622.0 + u \times 129.5$	622.0	129.5	76					
$200 + \exp(6.2 + u \times 0.3)$	715.4	158.2	48	0.31	0.003	0.019	0.020	0.003
$200 + \exp(6.0 + u \times 0.3)$	622.0	129.5	76					
$829.5 + u \times 193.2$	829.5	193.2	24	0.54	0.003	0.021	0.045	0.013
$622.0 + u \times 129.5$	622.0	129.5	38					
$200 + \exp(6.4 + u \times 0.3)$	829.5	193.2	24	0.54	0.001	0.009	0.028	0.007
$200 + \exp(6.0 + u \times 0.3)$	622.0	129.5	38					
$829.5 + u \times 193.2$	829.5	193.2	31	0.54	0.003	0.021	0.034	0.009
$622.0 + u \times 129.5$	622.0	129.5	31					
$200 + \exp(6.4 + u \times 0.3)$	829.5	193.2	31	0.54	0.003	0.015	0.023	0.005
$200 + \exp(6.0 + u \times 0.3)$	622.0	129.5	31					
Expected probabilities					0.005	0.025	0.025	0.005

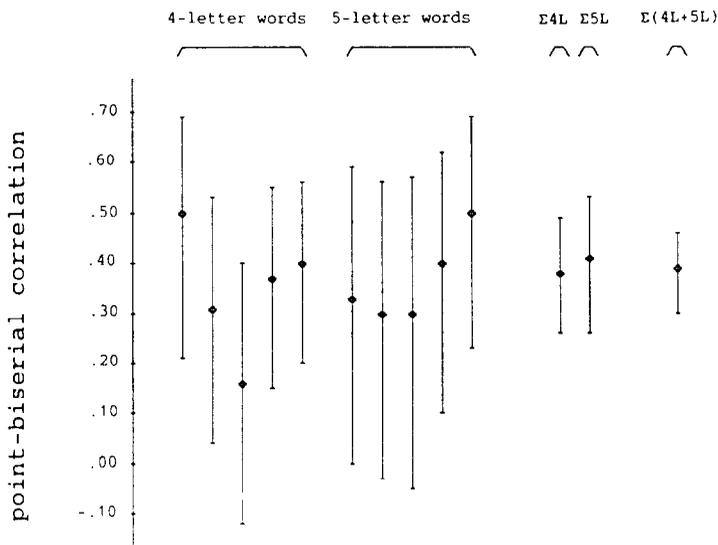


Fig. 6. Point biserial correlations and 95% confidence intervals for a VHF study in which five series of four-letter words and five series of five-letter words were administered (data from one subject). Data for the separate series, as well as for the sum of the four-letter word series ($\Sigma 4L$), the sum of the five-letter word series ($\Sigma 5L$), and the sum of all series ($\Sigma(4L + 5L)$).

$$r_{pb} = \text{sgn}(f(r_{pb})) \times \text{sqrt} \left\{ \frac{A \times (E-1)^2}{[D \times (E+1)]^2 - B \times (E-1)^2} \right\}, \tag{11}$$

$\text{sgn}(f(r_{pb})) = +1$ if $f(r_{pb}) \geq 0$ and -1 if $f(r_{pb}) < 0$,

$$D = \text{sqrt}(A + B),$$

$$E = \exp[f(r_{pb}) \times D \times 2].$$

Homogeneity of r point-biserial correlations is tested with the statistic of Equation (12), which has an approximate chi-square distribution with $r-1$ degrees of freedom (see the discussion of the lambda index for more information).

$$\sum u_i^2 = \sum [f(r_{pb_i}) - f(r_{pb_{est}})]^2 / s_i^2, \tag{12}$$

$$s_i^2 = 1/N_i,$$

$$f(r_{pb_{est}}) = [\sum f(r_{pb_i}) / s_i^2] / \sum 1/s_i^2.$$

If $\sum u_i^2$ is significant, contrasts can be tested using the same equations as Equation (6) and Table 3.

The distribution of $f(r_{pb_{est}})$ is approximately normal with variance $1/\sum(1/s_i^2)$, so that a significant difference from zero of the grand point-biserial correlation is easily testable. Table 5 gives the result of the analyses for homogeneity and significance of $f(r_{pb_{est}})$ for the data of the experiment leading to Table 2 [4].

Table 5. Point-biserial correlations of five replications of a series of four-letter words (Series 1-5), and of a series of five-letter words (Series 6-10). Grand r_{pb} calculated with the use of Equations (10), (11), and (12). probability of grand r_{pb} being equal to zero calculated with Equations (8), (10), and (12). probability of the 10 r_{pb} s being homogeneous calculated with Equations (10) and (12). Underlined r_{pb} s differ at a 5% level from each other [Equations (10) and Table 3].

	Series										Grand	P	P
	1	2	3	4	5	6	7	8	9	10	r_{pb}	(=0)	(homog)
s1	0.50	0.31	0.16	0.37	0.40	0.33	0.30	0.30	0.40	0.50	0.39	0.000	0.736
s2	0.24	0.09	0.07	0.21	0.22	0.25	0.35	0.42	0.60	0.40	0.31	0.000	0.026
s3	0.44	0.53	0.37	0.56	0.19	0.19	0.59	0.42	0.54	0.44	0.49	0.000	0.258
s4	0.29	0.35	0.31	0.34	0.35	0.20	0.31	0.00	0.23	0.38	0.30	0.000	0.589
s5	0.38	0.33	0.44	0.23	0.03	0.16	0.48	0.62	0.35	0.44	0.39	0.000	0.037
s6	0.27	0.25	0.24	0.30	0.07	-0.04	0.00	0.46	0.40	-0.06	0.20	0.001	0.458
s7	0.15	0.18	0.36	0.17	0.25	0.23	0.55	0.54	0.15	0.54	0.32	0.000	0.243
s8	0.27	0.36	0.12	0.36	0.34	0.08	0.38	0.35	0.17	0.55	0.31	0.000	0.768
s9	0.13	0.07	0.18	0.16	0.21	-0.75	0.29	0.22	0.12	0.23	0.17	0.002	0.549
s10	-0.03	0.34	0.25	0.12	0.22	0.29	0.35	0.21	0.19	0.32	0.23	0.000	0.755
s11	-0.11	-0.01	0.09	-0.11	0.00	0.33	0.07	0.23	0.11	-0.03	0.05	0.270	0.471
s12	-0.03	0.02	-0.01	0.19	0.17	0.14	-0.08	0.26	0.05	0.27	0.10	0.028	0.682
s13	-0.07	0.30	0.29	0.25	0.00	0.32	0.23	0.49	0.17	0.42	0.25	0.000	0.054
s14	0.32	0.17	0.25	0.20	0.15	-0.01	0.42	0.33	0.07	0.19	0.21	0.000	0.778

CONCLUDING REMARKS

The foregoing should provide any reader with the necessary information to make firm statements about the performance of individual subjects and/or patients. Graphical analyses allow to see whether the pattern of results is stable (see Fig. 1, 2, and 4). Statistical analyses

make it possible to assess the reliability of obtained VHF differences, both within and across experimental sessions, and to calculate confidence regions around the parametric values obtained (see Fig. 3 and 6). Though the calculations involved may look a little bit complicated to a not mathematically oriented investigator, they are very easy to implement in a user friendly computer program.*

A warning, however, must be given to conclude. No matter how stable and reliable the results of a subject may be, reliability alone never suffices to generalize from data to, for instance, statements about the cerebral organization of a subject. To do so, the results must first be compared to other measures of the concept under investigation. As an example, let's see what we can conclude from the data of the experiment displayed in Tables 2 and 5 (also see [4]). About 50 words were presented to the LVF and the RVF. Because processing of words is known to be a function of the dominant cerebral hemisphere, the data urge one to conclude that all subjects, except subject 11, have left hemisphere dominance, and that subject 11 has a bilateral cerebral organization. However, because words are found to favor the RVF for more reasons than left hemisphere dominance for verbal stimuli [3, 7, 11, 17], it might be asked whether the lack of a clear RVF advantage for subject 11, and to a lesser extent for subject 12, could be an indication of right hemisphere dominance. To check this possibility, we ran two further experiments on the two subjects.† Four replications of an object naming latency test (ONLT) and two replications of a clock reading latency test (CRLT) were administered. Prior experimentation had indicated that the ONLT leads to a RVF advantage in right-handed subjects (mean $r_{pb} = 0.06$, also see [15]), and that the CRLT leads to a LVF advantage in right-handed subjects (mean $r_{pb} = -0.06$, also see [1, 15]). The results of the four replications of ONLT for subject 11 were: 0.06, 0.03, 0.05, and 0.21; those for the two replications of the CRLT: -0.03 , and -0.06 . Thus, the overall picture of subject 11 contradicts the possibility of right hemisphere dominance; it even questions the possibility of bilaterality suggested by Tables 2 and 5. On the other hand, the results of subject 12 on the ONLT and the CRLT were respectively ONLT: -0.25 , -0.45 , 0.00 , -0.22 , and CRLT: 0.01 , 0.05 ; which is much stronger evidence for right hemisphere dominance. The results of the ONLT and the CRLT, therefore, seriously question the validity of a word VHF recognition task as the sole test of individual human cerebral asymmetry. We are now investigating the merit of a battery of three tests, consisting of word recognition, ONLT, and CRLT, in the hope that this will suffice to distinguish satisfactorily between left and right cerebral dominant subjects.

Acknowledgment—The authors wish to thank the two anonymous reviewers for their many helpful comments on an earlier draft of the manuscript.

REFERENCES

1. BERLUCCHI, G., BRIZZOLARA, D., MARZI, C. A., RIZZOLATTI, G. and UMILTÀ, C. The role of stimulus discriminability and verbal codability in hemisphere specialization for visuospatial stimuli. *Neuropsychologia* 17, 195-202, 1979.
2. BRYDEN, M. P. *Laterality: Functional Asymmetry in the Intact Brain*. Academic Press, New York, 1982.

*The programs written for this article that will run on most IBM compatible microcomputers, can be obtained by sending a 5.25 inch diskette in a returnable box to Marc Brysbaert, Department of Psychology, University of Leuven, B-3000 Leuven, Belgium. Ten U.S. dollars need to be enclosed for administration costs.

†These tests were administered several months after the word recognition test, because the results of the two subjects in subsequent experiments were too different (subject 12 showing more indications for right hemisphere dominance than subject 11).

3. BRYLSBAERT, M. and D'YDEWALLE, G. Callosal transmission in reading. In *Eye Movement Research: Physiological and Psychological Aspects*, G. LÜER, U. LASS and J. SHALLO-HOFFMANN (Editors), pp. 246-266. Hogrefe, Toronto, 1988.
4. BRYLSBAERT, M. and D'YDEWALLE, G. Tachistoscopic presentation of verbal stimuli for assessing cerebral dominance: Reliability data and some practical recommendations. *Neuropsychologia*, in press.
5. DAS GUPTA, S. Point biserial correlation coefficient and its generalization. *Psychom.* **25**, 393-408, 1960.
6. HELDIGE, J. B., BLOCH, M. I. and TAYLOR, A. K. Multitask investigation of individual differences in hemispheric asymmetry. *J. exp. Psychol.: HPP* **14**, 176-187, 1988.
7. HERON, W. Perception as a function of retinal locus and attention. *Am. J. Psychol.* **70**, 38-48, 1952.
8. HOLLAND, B. S. and COPENHAVER, M. D. Improved Bonferroni-type multiple testing procedures. *Psychol. Bull.* **104**, 145-149, 1988.
9. HOLM, S. A simple sequentially rejective multiple test procedure. *Scand. J. Statist* **6**, 65-70, 1979.
10. KAZDIN, A. E. Statistical analysis for single-case experimental designs. In *Single Case Experimental Designs* (2nd edn), H. BARLOW, and M. HERSEN (Editors), pp. 285-324. Pergamon, New York, 1984.
11. KIRSNER, K. and SCHWARTZ, S. Words and hemifields: do the hemispheres enjoy equal opportunity? *Brain Cognit.* **5**, 354-361, 1986.
12. LEVY, J. Individual differences in cerebral hemispheres asymmetry: Theoretical issues and experimental considerations, In *Cerebral Hemisphere Asymmetry: Method, Theory, Applications*, J. B. HELDIGE (Editor), pp. 465-497. Praeger, New York, 1983.
13. LUCE, R. D. *Response Times*. Oxford University Press, New York, 1986.
14. MARASCUILO, L. A. Large-sample multiple comparisons. *Psychol. Bull.* **65**, 280-290, 1966.
15. MCKEEVER, W. F. Tachistoscopic methods in neuropsychology. In *Experimental Techniques in Human Neuropsychology*, H. J. HANNAY (Editor), pp. 167-211. Oxford University Press, New York, 1986.
16. MILLER, J. A warning about median reaction time. *J. exp. Psychol.: HPP* **14**, 539-543, 1988.
17. O'REGAN, J. K. and LÉVY-SCHOEN, A. Eye-movement strategy and tactics in word recognition and reading. In *The Psychology of Reading*, M. COLTHEART (Editor), pp. 363-384. Lawrence Erlbaum, London, 1987.
18. SPROTT, D. A. and BRYDEN, M. P. Measurement of laterality effects. In *Cerebral Hemisphere Asymmetry: Method, Theory, and Applications*, J. B. HELDIGE (Editor), pp. 445-464. Praeger, New York, 1983.
19. YOUNG, A. W. and ELLIS, A. W. Different methods of lexical access for words presented in the left and right visual hemifield. *Brain Lang.* **24**, 326-358, 1985.