Pupillary dilation as a measure of attention: A quantitative system analysis

BERT HOEKS and WILLEM J. M. LEVELT
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

It has long been known that the pupil dilates as a consequence of attentional effort. But the function that relates attentional input to pupillary output has never been the subject of quantitative analysis. We present a system analysis of the pupillary response to attentional input. Attentional input is modeled as a string of *attentional pulses*. We show that the system is linear; the effects of input pulses on the pupillary response are additive. The impulse response has essentially a gamma distribution with two free parameters. These parameters are estimated; they are fairly constant over tasks and subjects. The paper presents a method of estimating the string of attentional input pulses, given some average pupillary output. The method involves the technique of deconvolution; it can be implemented with a public-domain software package, PUPIL.

The primary function of the pupillary reflex is to regulate the amount of light entering the eye, both in response to changes in the incident illumination (Lowenstein & Lowenfeld, 1962; Young & Biersdorf, 1954) and in order to maintain visual acuity under changes in the state of accommodation of the eye (Lowenstein & Lowenfeld, 1962). However, under conditions of constant illumination and accommodation, pupil size has been observed to vary systematically in relation to a variety of physiological and psychological factors, including nonvisual stimulation, habituation, fatigue, sexual and political preference, and level of mental effort (Goldwater, 1972; Tryon, 1975). All these sources of pupillary variation can be headed with the word *attention*.

Although dilation of the pupil in response to increased attention was first observed early in this century (Lowenstein, 1920), the first systematic study of the phenomenon appears to have been that of Hess and Polt (1964). In this study, subjects were required to mentally solve a series of multiplication problems varying in difficulty. Typically, what was observed in this task is that in the course of presentation of the problem and its solution by the subject the pupil would gradually dilate, reaching its maximum prior to the verbal report, and then return to its original size. It was also found that the more difficult the problem, the greater the degree of dilation. The usefulness of the pupillary response as an index of attentional

effort was further demonstrated in a series of elegant studies by Kahneman and his associates (see Kahneman, 1973). Their work, as well as subsequent research, has shown that pupil size can serve as an index of processing load in mental arithmetic tasks, language processing tasks, and short-term memory tasks, as well as in reaction time tasks in which stimulus probability is varied. See Goldwater (1972), Janisse (1977), and Beatty (1982) for reviews of this work.

Although these studies established the validity of the pupillary response as an indicator of attentional effort, they did not establish it as a measure in a stricter sense. In particular, the function relating attentional effort to the pupillary response was never analyzed. The present paper is an attempt to fill this gap. In it, we present a system analysis of the pupillary response to attentional effort. This, in turn, provides a method of computing the *attentional input*, given a measured *pupillary output*.

The paper will proceed as follows. We will first introduce the model, which relates attentional input to pupillary output. Next we will discuss how the model's free parameters can be estimated. The basic method is deconvolution, and we will apply it to a set of experimental data collected for this purpose. Third, we will outline how in practice the underlying attentional input can be computed, given our estimated parameters and a measured pupillary response. The paper will close with a discussion of the method's potential and limitations.

THE MODEL

Input and Output

The model relates attentional input to pupillary output. How does one model attentional input? Most tasks are attentionally complex. Even a simple Donders reaction task involves a variety of attentional responses on the subject's part. There is the expectancy response when the stimuli are equally spaced; there is the perceptual response to the

appearance of the stimulus; there is the decision to respond; and there is the initiation of the finger press. These attentional responses come at different moments in time, and they can be of different magnitudes. Correspondingly, we have chosen to model attentional input as a *sequence of attentional pulses* that can vary in number, in temporal distribution, and in pulse amplitudes. The total attentional effort involved in a response is the sum of amplitudes of the attentional pulse train.

Pupillary output is a less abstract entity. It is the continuously varying deviation of the pupil's diameter from the baseline value. The latter is the value just prior to the stimulus (task, instruction, event, whatever) that initiates the attentional response. In practical measurement, this response is discrete, a string of diameter values for discrete time intervals of 20 msec. These time intervals can be numbered from 1 to t, where t is the moment the pupil returns to baseline without further significant deviations. Hence, the output can be represented as a vector in t-dimensional space, with successive time intervals as dimensions and pupillary deviations from baseline as values. We will call this the T-space.

Linearity

In the model, the system characteristics are taken to be constant during a measurement session. This means that the same attentional input pulse always generates the same pupillary response. In addition, the output is assumed to obey the superposition principle: suppose there are two input pulses or input pulse trains $x_1(t)$ and $x_2(t)$, with corresponding outputs $y_1(t)$ and $y_2(t)$. The output in response to a new input $x_1(t)+x_2(t)$ is $y_1(t)+y_2(t)$. This is equivalent to saying that input and output are related in a linear way, or that the system is linear. These assumptions are graphically shown in Figure 1.

Given these assumptions, one can relate the system's input and output thus:

$$y(t) = h(t) * x(t), \tag{1}$$

where $y(t)$ is the output, $x(t)$ is the input, and $h(t)$ is the impulse response of the system. The $*$ is the "convolution operator."

The Impulse Response

The impulse response is the system characteristic that is constant over time. In order to derive $h(t)$, a more detailed model must be developed. This is necessary, because the complexity of the pupil's response requires its reduction to a sequence of more elementary neurologically based processes. We propose a cascade model, with a number of layers or *boxes*, with information flowing from layer to layer, or from box to box. Each layer in the model has its own impulse response. We assume that for each layer this impulse response is a declining exponential function.

$$h_i(t) = b_i e^{-a_i(t-t_{0,i})} \quad t > t_{0,i}$$
$$h_i(t) = 0 \quad t \leq t_{0,i} \tag{2}$$

where $h_i(t)$ is the impulse response of box i and a_i, $t_{0,i}$ and b_i are positive constants.

Given this cascade of elementary responses, the impulse response of the system as a whole will have the form of a general gamma function. Its parameters are to be estimated from experimental data. But because the general gamma function has as many describing parameters as there are layers or successive boxes in the model, we will, in general, not be able to derive unique or stable estimates of these parameters. Hence we make the additional assumption that $a_i = a_j$ for all i and j; that is, the impulse responses of all layers are the same except for an amplification factor. Under these conditions the general gamma function reduces to the Erlang gamma function:

$$h_{tot}(t) = t^n e^{-nt/t_{max}} \quad t > 0$$
$$h_{tot}(t) = 0 \quad t \leq 0 \tag{3}$$

where $n+1$ is the number of layers and t_{max} is the position of the response's maximum. The parameters n and t_{max} fix the form of the Erlang gamma function.

The Output in Terms of Pupil Size

As we stated, the output of the model is the pupil diameter's continuing deviation from baseline. It is, however, not self-evident that the assumption of linearity will stand an empirical test when the output is measured in straight pupil diameter values. In fact, one could argue that it is the *area* of the pupillary change that matters—that is, the squared change in diameter. Which exponent is correct? We argue that it is immaterial which exponent is taken. When the system is linear (as we hope), it will be linear for any exponent, and hence for $m = 1$.

Assume that the pupil starts in a resting state and that an attentional pulse will cause a linear change in the *area* of the pupil. Then, pulse amplitude A will relate to the difference between the old and new pupil areas as follows (where $\vdash$ stands for *related*):

$$A \vdash Area_{new} - Area_{old}$$

$$A \vdash d_{new}^2 - d_{old}^2$$

if the new pupil diameter $d_{new} = d_{old} + \partial d$ and ∂d is small, then

$$A \vdash 2d_{old}\partial d + \partial d^2 \sim 2d_{old}\partial p. \tag{4}$$

Equation 4 shows the linear relation between pulse amplitude A and the change in pupil diameter ∂d if the pupil exponent is two and the pupil starts in a rest state—that is, at base level. But Equation 4 can be generalized to any exponent k. For any k and small pupillary changes, Equation 5 will hold:

$$A \vdash k d_{old}^{k-1} \partial d. \tag{5}$$

It follows from Equation 5 that the "real" power of the pupil is quite immaterial for our procedure. For any exponent k, the changes of the pupil's diameter will show a linear relation to the amplitude of attentional pulses, as long as ∂d is small. *But this holds only when our linearity*
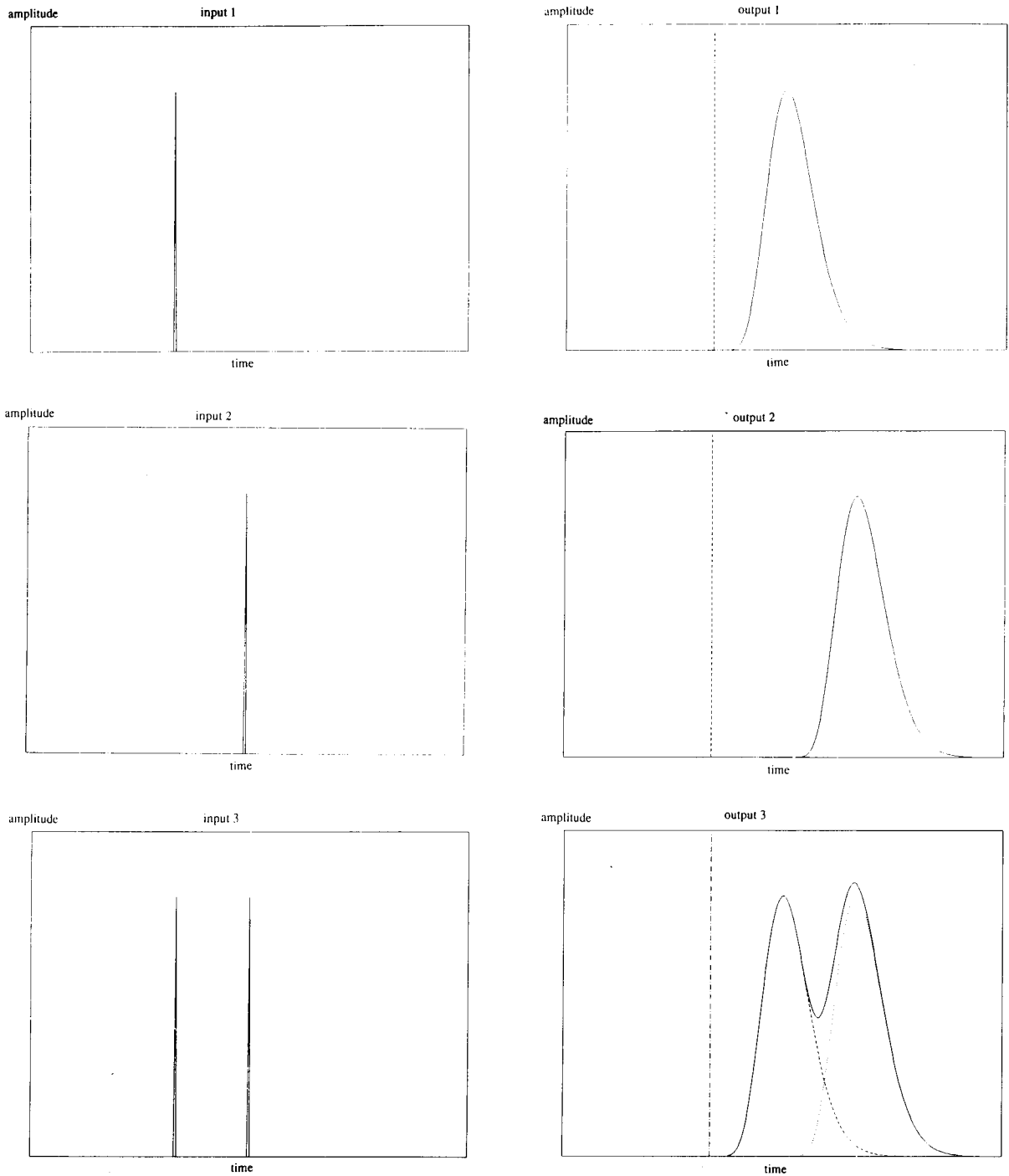
Figure 1. Convolution. The shape of the input is the same in Situations 1 and 2, but in Situation 1 the event takes place at an earlier moment than it does in Situation 2. The corresponding Outputs 1 and 2 are identical in form, but Output 2 is moved in time. Input 3 is a concatenation of Input 1 and Input 2. Correspondingly, Output 3 is the sum of Output 1 and Output 2.

assumption holds for the system. In other words, one way of verifying that assumption is to show that the pupil diameter's change relates linearly to attentional input, $m = 1$.

ESTIMATING THE MODEL'S PARAMETERS

In order to make the model work—that is, to use it for the computation of attentional input, given some pupillary output—three parameters must be determined. For this, we will use two different deconvolution methods.

The first parameter to be estimated is the pupil diameter's exponent for which the system is linear. As outlined in the previous section, we expect this exponent to be 1. Here we will use the filter of Bracewell and Helstrom (see Jansson, 1984) with some adaptations.

The second and third parameters are the two free parameters of the Erlang gamma distribution, n and t_{max}, that is, the number of boxes in the cascade $(n+1)$ and the position of the response maximum. Here we will supplement Bracewell and Helstrom's method (Jansson, 1984) with a least squares estimation procedure. Both methods are described in the Appendix.

These estimations must be based on relevant empirical data, and we will shortly describe the experimental procedure used to obtain them. It was our hope that the three parameters would be sufficiently stable over subjects and tasks. In that case, there would be no real need to estimate them for each new subject or experiment.

We proceed now as follows. The experimental procedure will be introduced first. We will then turn to estimating the pupil diameter's exponent. If it is close to 1, we will have good evidence that the system is indeed linear. Finally, the gamma distribution's parameters n and t_{max} will be estimated.

Method

The aim of the experiment was to collect a range of pupillary responses that would allow us to verify the system's linearity and to estimate the system's parameters. In addition, we wanted to obtain evidence about the stability of these parameters over subjects and tasks. Basically the experiment consisted in measuring 8 subjects' pupillary reactions in a simple reaction task. They were presented with an acoustic or visual stimulus to which they had to respond. In one condition, the response was a push-button reaction; in another condition, it was merely the subjects' internal reaction—no overt response was required.

Apparatus. The subjects were tested individually in a laboratory room that contained the complete Whittaker 1998-S Eye View Monitor and TV-pupillometer System, and its computer monitor, both connected to a PDP-11/73 system. The illumination of the room was normal, from strip lighting. Each subject was seated in an adjustable chair with back headrests. During the experimental runs, the subject viewed a fixation point on a monitor at a distance of approximately 1 m at eye level, while a video camera monitored the subject's left eye. In this way, reflections were recorded from an infrared source light that was directed continuously to the eye. Every 20 msec, the pupil diameter was automatically measured as the number of scan lines that intersected the image of the pupil on the experimenter's monitoring screen. The spatial resolution is about 0.05 mm. The experiment was controlled by a set of computer programs. One program generated tones that were presented to the subject through headphones. Another program presented pictures to the subject on a video screen. The intensity of the screen was set at a low level so that the picture was just visible with the room's strip lighting on.

Stimuli. The auditory stimuli were 1000-Hz tones at a convenient loudness level, lasting 100 msec. They came from a loudspeaker in front of the subject. The visual stimuli were white outline circles on a constant gray background. They were displayed with a radius of 2.0 cm around the fixation point on the screen and lasted 100 msec. The luminance of the stimulus was $(6.24 \pm 0.03)/10$ cd/m² and the background had a luminance of $(6.17 \pm 0.03)/10$ cd/m² (measured 1 m from the video screen with an 80× optometer from United Detector Technology).

To verify the superposition principle (see the section above on linearity), we needed responses to single stimuli and to stimuli in relatively close succession. Three kinds of trials were used:

Singleton trials. There was only one stimulus.

Close pair trials. There were two stimuli in close succession, with 640 msec between stimulus onsets. The subject had to react to both stimuli.

Distant pair trials. These were the same as the close pair trials, but the stimuli were 1,640 msec apart.

The interval between two trials (i.e., between the onset of the first stimulus in a trial and the onset of the first stimulus in the next trial) varied randomly between 5.0 and 6.0 sec. A session contained 150 trials, 50 of each kind, in random order. Within each session, all stimuli were either auditory or visual.

Subjects. Eight students (4 males and 4 females) were paid to participate in the experiment. All subjects were naive with respect to the experimental task.

Procedure. The experiment was divided into four sessions. Each session lasted about 15 min. In two sessions, the stimuli were auditory, and in the other two, they were visual. For each kind of stimulus, there was one session in which subjects had to press the push button as fast as possible every time he/she perceived a stimulus. In the other two sessions, no push-button response was required. Prior to each session, a subject read an instruction. After that, during an interval of 1 min, stimuli were presented for exercise. Then the subject could ask questions, after which the test session began.

The order of the four sessions was varied systematically over the subjects, with the restriction that each subject began with either two visual stimulus sessions or two auditory stimulus sessions. Between the second and the third sessions, the subject took a coffee break. The experiment took about 1.5 h for each subject. All sessions were run in the afternoon.

Results

The pupil responses on all 50 trials of a single kind were sampled and averaged. In this way, three averaged pupil traces were calculated for each subject and session, for singletons, for close pairs, and for distant pairs.

The pupillary responses in the sessions where no push-button response was required were too small and noisy for further data analysis. Therefore we decided to limit the analyses to the data from the other two sessions.

Determining the pupillary exponent. To determine the exponent for which the pupil diameter has linear behavior, the output $y_m(t)$ was calculated:

$$y_m(t) = C_m [\theta(\partial d)]^m, \tag{6}$$

where C_m is a constant such that the maximal value of $y_m(t)$ is 1, $\theta(\ldots)$ is the Heaviside function

$$\theta(x) = 0 \qquad x < 0$$
$$\theta(x) = x \qquad x \geq 0,$$

∂d is the pupil's dilation with respect to the baseline (see Equation 5), and m is the exponent to be estimated.

Because the exponent m can only be determined by assuming the superposition principle, we used the close pair trials for its computation. Only in this case would pupillary responses sufficiently overlap. Assuming that each stimulus in the pair generates the same attentional response, the output is a linear combination of two identical pupil responses, shifted in time.

The deconvolution method, described in the Appendix, allows us to compute an optimal estimation of these two identical responses that are spaced 640 msec apart. If the system is linear, their sum should be a close approximation of the measured pupillary response. The method was applied for values of m ranging from 0.5 to 3.0, first in steps of 0.1, later around the optimum in steps down to 0.001.

To evaluate the results, a measure of approximation between computed and measured pupillary responses was developed. Both the measured and the computed pupillary responses can be represented as vectors in T-space (see the section on Input and Output above); let them be called $f(t)$ and $g(t)$ for the measured and computed values, respectively. If the approximation is perfect, the two vectors will coincide. If not, there will be some nonzero angle τ between them. The size of this angle is an inverse measure of fit. In more detail, the inproduct of the two vectors can be defined as

$$\langle f,g \rangle = \sum_k f(t_k)g(t_k), \tag{7}$$

where k ranges from 1 to t.

The norm or size of f is then $|f| = \sqrt{\langle f,f \rangle}$. The angle τ between $f(t)$ and $g(t)$ can now be defined as

$$\cos \tau = \frac{\langle f,g \rangle}{\sqrt{(|f||g|)}}. \tag{8}$$

If we assume that $|f| \sim |g|$, Equation 11 becomes

$$\cos \tau = \frac{\langle f,g \rangle}{|f|} . \tag{9}$$

With τ as our measure of fit, the optimal m was determined for each subject and stimulus mode (auditory vs. visual). Here are the results: For the auditory trials, we found $m = 1.22$ with a standard deviation of .63 over the 8 subjects. For the visual trials, these values were $m = 1.27$ and $\sigma = .63$. Neither of these m values differed significantly from 1; also they did not differ significantly from each other. This is in full agreement with the expectations formulated above. In addition, the average value of τ was as small as $1.29°$, which means that the measured output was reproduced well by the computations. In fact, $\tau = 1.29°$ indicates that 97.75% of the variance in the data was reproduced by the computations. This is strong evidence for the validity of our linearity assumption.

Estimating n and t_{max}. The parameters that describe the form of the impulse response are the number of layers (or boxes), $n+1$, and the position of the response maxi-

mum, t_{max}. To estimate these parameters, we used all three kinds of trials (singletons and both the close and distant pairs) for both stimulus modes. Again, the procedure consisted of deconvolution (see the Appendix). Our model's assumption is that the input is a string of attentional pulses. Deconvolution is a method of computing an input string that reproduces the output, given a particular impulse response (i.e., with specific n and t_{max}). Optimal values for n and t_{max} can be found by applying deconvolution to a wide range of n and t_{max} values. The pair of values for which the output is best reproduced is the optimum. The goodness of fit will again be determined by τ, the angle between the computed and measured vectors in T-space.

However, there is an additional degree of freedom, the hypothetical number of attentional pulses. The approximation of the pupillary response will, of course, improve with the number of attentional pulses that we allow as input. We decided to compute optimal estimations for n and t_{max} for any number of pulses between 1 and 7. Angle τ will decrease with the number of pulses. We will then go for the smallest number of pulses beyond which there is no substantial improvement of fit anymore.

How does one compute the parameters n and t_{max} for a given number of input pulses? We used two factorial designs and then a single-step optimizing procedure. In the first factorial design, we varied t_{max} between 0.7 and 1.2 sec with a step width of 0.1 sec, and n between 4 and 14 with a step width of 2.[1] Next, we carried out a second factorial design in which we used the optimal parameters found. Here n and t_{max} were varied with a step width of 0.04 sec for t_{max} and 0.8 for n. The single-step method started with the optimal parameters found in the second factorial design. The first step width for n was 0.4 and for t_{max} 0.02 sec. For the best parameters, the estimated output for $n \pm 0.4$ and $t_{max} \pm 0.02$ were calculated. The best parameters were again chosen and the step width was divided by 2. This procedure was carried out five times, eventually leading to the final solution.[2] The parameter estimations are presented in Table 1.

Before discussing these results, we should say a word about the number of input pulses we allowed. As mentioned, it should be the smallest number beyond which the fit does not improve substantially. How far should τ decrease for us to accept the solution? We settled for $\tau = 5.5°$. The reason can be seen from Table 1, which also represents the τ values. There are 48 (n,t_{max}) estimates in the table. With the τ criterion set at $5.5°$, in 13 cases no solution could be found with 7 input pulses or less. Even a small decrease of our criterion, to $5.0°$, dramatically increases the number of nonsolutions to 23. This is about one half the cases, showing that our criterion has become too stringent. It should be added that a fit of $5.5°$ corresponds to a quite satisfactory 91% explained variance. Figure 2 shows the two best and the two worst solutions within the $5.5°$ criterion. They show the measured output, and its computed approximation, as well as the pattern of attentional input pulses.

Table 1
Estimated Impulse Response Parameters n and t_{max} for Auditory (A) and Visual (V) Singletons, Close Pairs, and Distant Pairs

Subject	Mode	n	t_{max}	No.	τ
Singletons					
1	A				>5.5
	V				>5.5
2	A				>5.5
	V	12.9	0.68	6	5.33
3	A	16.4	0.98	2	4.77
	V	5.2	0.66	2	2.60
4	A	13.7	1.00	4	4.64
	V	11.6	0.82	3	4.21
5	A	11.4	0.71	3	4.98
	V	9.6	0.63	5	3.81
6	A	4.5	1.30	2	5.49
	V	3.8	1.15	3	4.49
7	A	11.1	1.12	2	4.46
	V	14.7	0.58	5	5.30
8	A				>5.5
	V	4.4	1.06	2	4.34
Close Pairs					
1	A	5.3	0.76	6	4.08
	V	12.9	0.77	5	5.26
2	A	16.5	0.97	4	5.37
	V	13.6	0.90	5	4.96
3	A	5.1	0.66	2	3.89
	V	12.2	1.02	3	3.87
4	A	3.7	0.95	4	4.93
	V	13.4	1.16	3	4.02
5	A	12.5	0.76	7	4.77
	V				>5.5
6	A	5.3	1.21	2	3.99
	V	6.8	1.14	2	5.12
7	A	11.0	0.91	3	2.81
	V				>5.5
8	A				>5.5
	V				>5.5
Distant Pairs					
1	A	10.2	0.82	7	4.17
	V	11.6	0.82	6	4.95
2	A	9.7	0.71	5	5.38
	V				>5.5
3	A	8.8	0.88	7	4.99
	V	10.0	0.86	4	4.29
4	A				>5.5
	V	13.4	1.15	5	5.32
5	A	15.7	1.10	5	5.20
	V				>5.5
6	A	4.5	1.23	4	4.35
	V	3.6	0.94	2	4.93
7	A	15.7	0.94	6	5.42
	V				>5.5
8	A				>5.5
	V	12.8	0.98	5	4.48

Note—For each pair of parameters, the number of attentional pulses and the degree of fit (τ) are also given.

Table 1 shows the parameter estimates for all 8 subjects, stimulus modes (auditory, visual), and trial types (singletons, close pairs, distant pairs). We found an average t_{max} of 930 msec ($\sigma = 190$ msec) and an average n of 10.1 ($\sigma = 4.1$). Using t tests, we found no significant parameter differences between the three trial types. Also, there were no significant differences between stimulus modes. These effects were very strong: the largest t value was 0.75. This means that these were nonsignificant even if the number of subjects would have been six times as large. Finally, we could not reject the null hypothesis that subjects had identical impulse responses within singletons, close pairs, or distant pairs. Of the 21 paired comparisons between subjects, 0 pairs differed significantly (at the 5% level) for t_{max}, and 3 pairs for n. These findings support the notion that the system's impulse response is a constant over tasks and stimulus modes. Although the subjects did not differ significantly, one should recognize that the between-subject variance in n was relatively high. We will return to this point in the Discussion section.

Summarizing, given the parameter estimates, the average impulse response can be expressed thus:

$$h_{mean}(t) = t^{10.1}e^{-10.1t/930} \qquad (10)$$

HOW TO MEASURE ATTENTIONAL INPUT IN PRACTICE

Now that we know the system's impulse response, it is possible to apply the method in practice. Given some average pupillary response, a string of attentional input pulses (as in Figure 2) can be computed. For this, one can use the program PUPIL, a VAX FORTRAN program. PUPIL is in the public domain; it can be supplied at cost.

PUPIL's input consists of an average pupillary response. The average can be within a subject (i.e., there are repeated measures) or over subjects. It makes little sense to use PUPIL on single-trial data. The trial-to-trial variability of the pupillary response is too large for that. It is important that input data be cleaned up. In particular, they should be free of eyeblinks. The input consists of a string of pupil sizes starting at the resting state and returning to the resting level. Negative (i.e., values below resting level) should be corrected to 0. The maximal signal duration that PUPIL can handle is 40 sec.

In addition, PUPIL has a set of input options. The first is the maximal number of input pulses one is prepared to accept; another one is τ, the fit criterion. By default, PUPIL uses the impulse response specified in Equation 10 above, but it is possible to opt for other t_{max} and n values.

PUPIL's output is a file that consists of measured and estimated pupillary output, as well as the string of input pulses generating that output (as in Figure 2 above; notice that these pulses are narrow, but that they do have nonzero width). These data, as well as the corresponding τ values, are also made available numerically. Finally, PUPIL generates a measure of *total processing load* for any input data; this is the area under the output curve (which is linearly related to the area under the input pulses).[3]

DISCUSSION

Four points need further discussion: the system's delay, the variability of n, the system's stability within a session, and the method's temporal resolution.
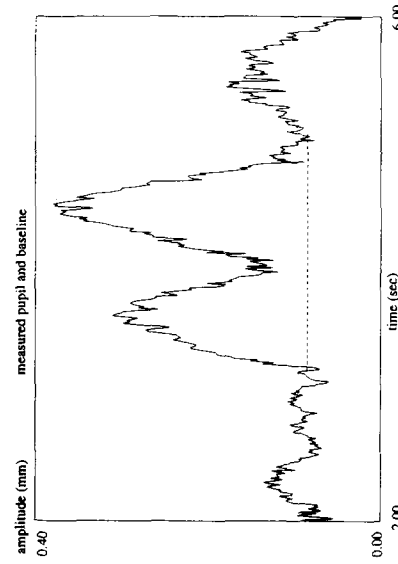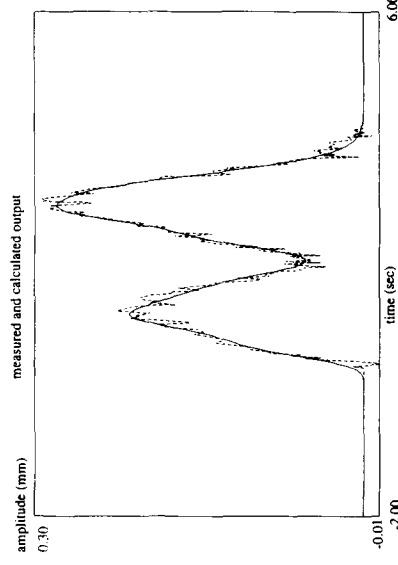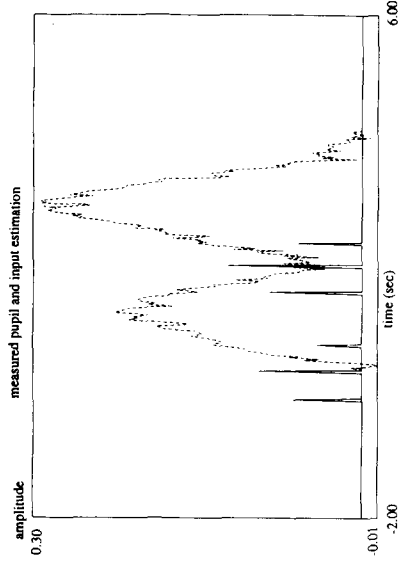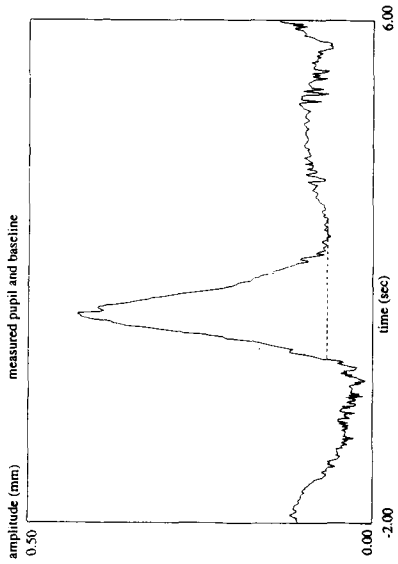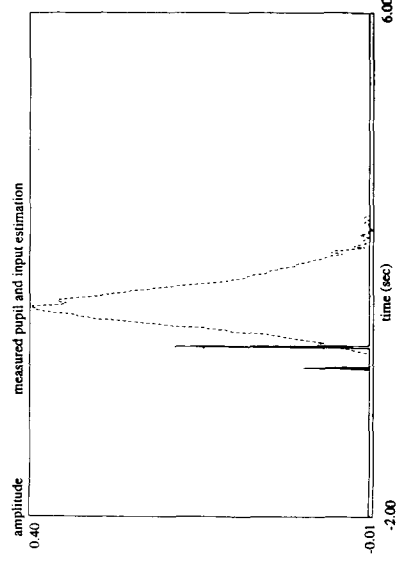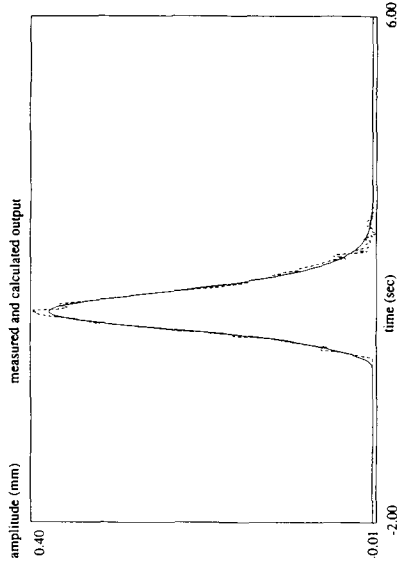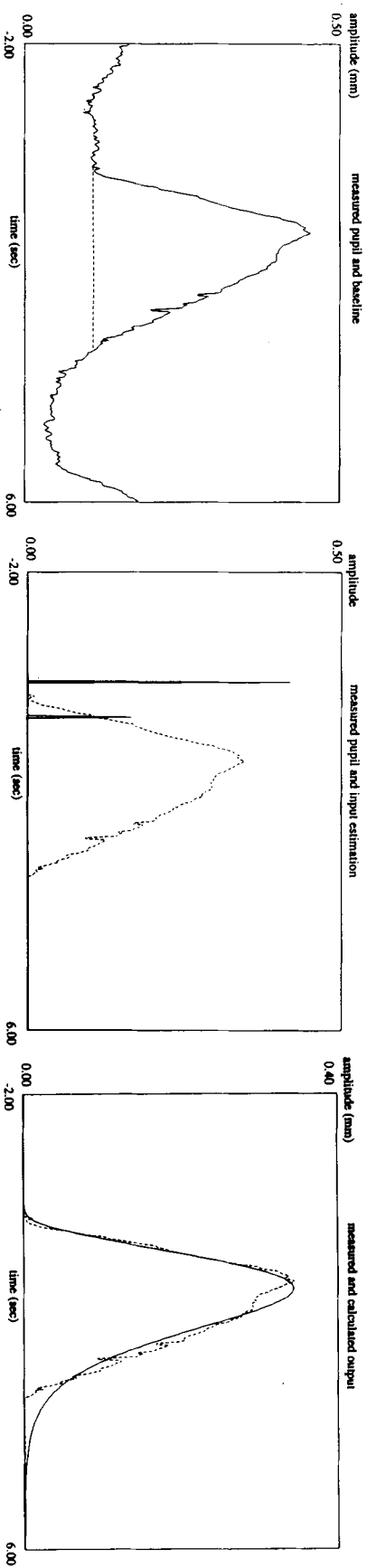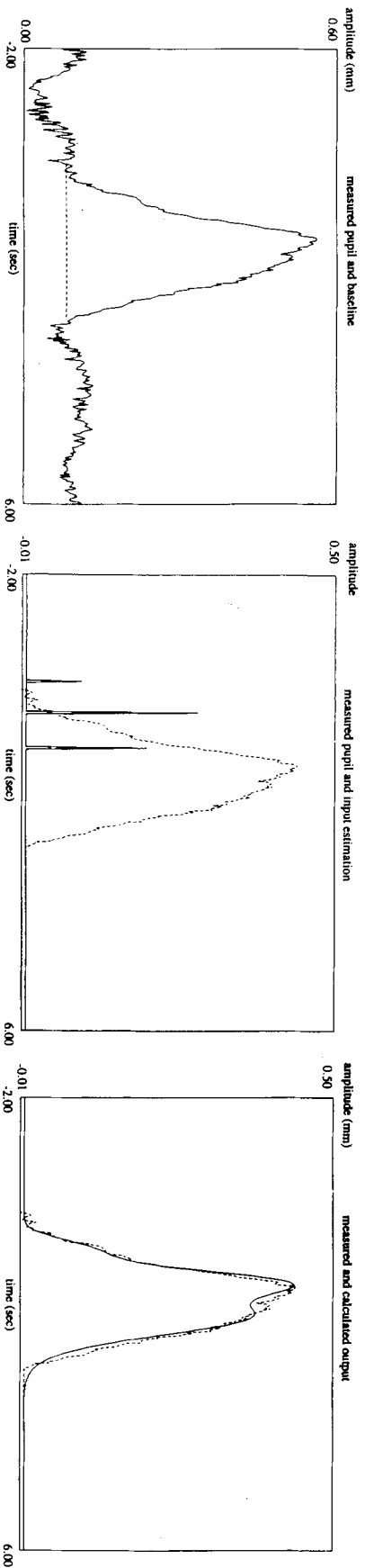
Figure 2. The two best (Subject 3, singleton, visual, and Subject 7, distant pair, auditory) and two worst cases (Subject 6, singleton, auditory, and Subject 7, close pair, auditory) within the $\tau = 5.5°$ criterion from Table 1. For each case, the left panel shows the averaged output and the baseline (broken line). The middle panel shows the measured output and the computed attentional peaks without delay. The right panel gives an impression of the goodness of fit, showing both the calculated and the measured outputs.

The System's Delay

If the subject's task is triggered by some stimulus at t_0, as in our experiment, the deconvolution method will produce a string of attentional pulses at moments t_i, t_j, and so forth. Is t_i really the moment of the first attentional pulse? No, it is not. Equation 2 does not involve a delay component. Deconvolution will put the first pulse exactly there where the pupil starts dilating and that is way too late. How much? This is a hard question. Data by Beatty (1982) and Zimmer (1984) show that the initiation of pupil dilation follows the stimulus at about 300–500 msec. In our data, the average first dilation came at 320 msec. But, of course, the first attentional pulse does not coincide with the stimulus either; it is a response *to* the stimulus. If its latency is l msec, then the system's delay is $320 - l$ msec. Can l be determined? At one level the answer is no. The attentional pulses are a theoretical construct that only exists in our model. They cannot be spotted in the subject's response. At another level, one might try to give them empirical content by relating them to event-related-potential components. For instance, Näätänen and Picton (1987) interpret the N100 wave as the first "cognitive" response to a stimulus. If we interpret that component as the first attentional pulse, then $l = 100$ msec or thereabouts. The system's delay is then something like 220 msec. This would mean that the computed pulse moments should all be decreased by 220 msec in order to find the "real" locations of the attentional pulses.

In practice, the delay problem will be negligible in most cases. The experimenter will usually be interested in the comparison between stimulus or task conditions. Do they differ in the size, number, or distribution of attentional pulses? Since the system's delay can be assumed to be a constant, the delay factor can be ignored in answering this type of question.

The Variability of n

Parameter n was estimated to be 10.1 ± 4.1. This is a large range, and one should wonder whether it is justified to use the default value $n = 10.1$ for all subjects. Luckily, deconvolution is quite insensitive to variations of n. This could be shown in a simulation where we took as input two attentional pulses, one at $t = 0$ and another one at $t = 600$ msec. We set t_{max} to 0.9 sec. When we varied n between 6 and 15 and applied deconvolution with $n = 9$ to the output curves obtained, the position of the two pulses ranged over no more than 60 msec, with a mean relative deviation in amplitude of less than 15%. In short, the method is quite robust with respect to variation in n.

Fluctuation of the Impulse Response Within a Session

Each of a subject's four sessions lasted about 15 min. Does a subject's impulse response show a systematic variation over a session? In order to test this, we split each session in two halves of 7.5 min. For each separate part,

we estimated the optimal impulse response for the three trial conditions. The three (n,t_{max}) estimations were then averaged. For the first 7.5 min, $(n,t_{max}) = 9.5 \pm 4.6$, 920 ± 230 msec. The result for the second part was $(n,t_{max}) = 8.8 \pm 4.1$, 850 ± 210 msec. For both n and t_{max}, it was not possible to reject the null hypothesis that the impulse responses were identical [for n, $t(59) = 0.47$, $p > .50$, and for t_{max}, $t(59) = 0.84$, $p > .20$]. Hence, we were justified in assuming that the impulse response was constant over the 15 min of measurement.

Temporal Resolution

What is the temporal resolution of the method? This question has two aspects. The first relates to the reliability of computed pulses' positions in time. The second relates to the discriminability of pulses. We are optimistic about the first aspect, less so about the second. In one simulation, we spaced three input pulses rather far apart, at $t = 0$, 600, and 1,200 msec. We computed the system's response, and then added 1%, 5%, or 10% noise. Deconvolution of these noisy data produced deviations from the input values of 0, 20, and 40 msec, respectively (there was a discrete 20-msec timescale, corresponding to the time grain of the Eye View Monitor). This showed that the computed pulse locations were quite reliable: the largest deviation was 40 msec, which is 3.3% of the entire range of the attentional pulses.

But how discriminable are attentional input pulses, or what is the resolving power of our deconvolution method? This we investigated in the following simulation. We used two-pulse inputs, and varied the pulse-to-pulse time interval. Again using the impulse response given in Equation 10, we computed outputs for each pulse pair. We then added 5% noise to these outputs, and applied deconvolution. It turned out that, for such noisy data, the method could no longer discriminate between pulses with a temporal distance smaller than 300 msec; it would then compute a single broad pulse instead of two narrow ones. It is important to keep a limit of this order in mind when one is interpreting experimental results.

This, then, brings us to a final theoretical question: Should one *want* to make finer discriminations? Our model assumes that there is attentional input of a string of pulses. Each pulse has infinitesimal duration. This is, of course, an idealization. All cognitive activities have *some* duration. Typical durations of event-related-potential components, some of which are psychologically interpretable, are between 100 and 300 msec. If these reflect real attentional waves (as one might suppose about a component such as P300, the odd-ball effect), the resolving power of our method, though not brilliant, is acceptable. One should, however, not expect that the method can be essentially improved. This is because the impulse response (Equation 10) essentially acts as a low-pass filter. The pupil simply does not transmit high-frequency components. Hence, they cannot be reconstructed by whatever deconvolution procedure.

REFERENCES

BEATTY, J. (1982). Task-evoked pupillary responses, processing load and the structure of processing resources. *Psychological Bulletin*, **91**, 276-292.

GOLDWATER, B. C. (1972). Psychological significance of pupillary movements. *Psychological Bulletin*, **77**, 340-355.

HESS, E. H., & POLT, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, **143**, 1190-1192.

JANISSE, M. P. (1977). *Pupillometry: The psychology of the pupillary response*. New York: Wiley.

JANSSON, P. A. (1984). *Deconvolution: With applications in spectroscopy*. Orlando, FL: Academic Press.

KAHNEMAN, D. (1973). *Attention and effort*. New York: Prentice-Hall.

LOWENSTEIN, O. (1920). Experimentelle Beiträge zur Lehre von den Katatonischen Pupillenveränderungen. *Monatschrift für Psychiatrie und Neurologie*, **47**, 194-215.

LOWENSTEIN, O., & LOWENFELD, I. E. (1962). The pupil. In H. Davson (Ed.), *The eye: Vol. 3. Muscular mechanisms* (pp. 256-329). New York: Academic Press.

NÄÄTÄNEN, R., & PICTON, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, **24**, 375-425.

TRYON, W. W. (1975). Pupillometry: A survey of sources of variation. *Psychophysiology*, **12**, 90-93.

YOUNG, F. A., & BIERSDORF, W. R. (1954). Pupillary contraction and dilation in light and darkness. *Journal of Comparative & Physiological Psychology*, **47**, 264-268.

ZIMMER, K. (1984). Ansätze der psychophysiologischen Indikation von Wissensrepräsentationen: Die Pupillomotorik als sensibler Indikator semantischer Informationsverarbeitungsaktivität. In F. Klix (Ed.), *Gedächtnis, Wissen, Wissensnutzung* (pp. 137-155). Berlin: VEB Deutscher Verlag der Wissenschaft.

NOTES

1. Values of t_{max} smaller than 0.7 sec too frequently led to solutions where the number of input pulses had to exceed 7 to reach a reasonable fit. Values beyond 1.2 sec are unrealistic, because in our measurements the maximal pupil size was reached no later than 1.3 sec after the stimulus. Given a standard delay in the pupillary system of at least 100 msec (see the Discussion), t_{max} values > 1.2 are unrealistic.

2. We used one additional optimizing rule that could protect us against continuing search in unrealistic areas of the (n, t_{max}) space. Suppose the system has a certain known impulse response. In that case, the least squares method gives a good estimation of the input—a series of narrow peaks or pulses. The filter of Bracewell and Helstrom, however, will calculate an input with relatively broad peaks. As a consequence, the Bracewell and Helstrom filter (see Jansson, 1984), different from the least squares method, estimates input peaks that are too broad. In case the system's impulse response is not known (i.e., as in our measurement situation), it may happen that the Bracewell and Helstrom filter occasionally produces a better output estimation than does the least squares method. This, however, can only be due to an erroneous choice of the impulse response's parameters. If they are correct, the least squares method should produce the better approximation. This lead us to use the following additional rule: Only those n and t_{max} are used for which the least squares method yields a better output estimation than does the filter of Bracewell and Helstrom.

3. The PUPIL package is available on the FTP account VMCMS.URC.KUN.NL (I.P. number 131.174.82.160), under user ANONYMOUS, password ANONYMOUS. The files are stored in the directory PUPIL.

APPENDIX

Deconvolution, More in Detail

Because the output is the convolution of the input and the impulse response, the inverse technique of convolution, deconvo-lution, calculates the input from the output and the impulse response. If we know the input, the impulse response can be calculated. We shall use two deconvolution methods. Before we describe them, the filter of Bracewell and Helstrom must be introduced (see Jansson, 1984).

The *convolution* technique can be described in the (usual) time domain:

$$y(t_n) = \sum_k h(t_{n-k}) x(t_k), \tag{A1}$$

where t_n is a discrete time moment ($t_n = t_0 + n\Delta t$); $y(t_n)$, the output; $h(t_n)$, the impulse response; and $x(t_n)$, the input in the time domain.

The output $y(t_n)$ can be transformed to the frequency domain by using a fast Fourier transform algorithm. Equation A1 is then replaced by the following product:

$$Y(f) = H(f)X(f), \tag{A2}$$

where f is the frequency; $Y(f)$, the output; $H(f)$, the impulse response; and $X(f)$, the input in the frequency domain.

If we want to calculate the input, we divide $Y(f)$ by $H(f)$. This estimate of the input is not very stable. To use a fast Fourier transform, $H(f)$ must have some special properties. One of these properties is that $H(f)$ ''becomes small'' for increasing f:

$$\lim_{f \to \infty} H(f) = 0. \tag{A3}$$

This means that $|1/H(f)| \to \infty$ if $f \to \infty$. If noise in the signal changes $Y(f)$ just a little, the input estimate becomes unstable. For this reason, Bracewell and Helstrom developed a filter which obviates this problem:

$$B(f) = H^c(f)/(|H(f)|^2 + N_0) \tag{A4}$$

$$X_{est}(f) = B(f)Y(f) \tag{A5}$$

where $B(f)$ is the filter of Bracewell and Helstrom, $X_{est}(f)$ is the input estimate in the frequency domain, $H^c(f)$ is the complex conjugated of the impulse response in the frequency domain, and N_0 is a positive constant.

It can be shown that if the noise is additive and has a Gaussian distribution, there exists no better filter than Bracewell and Helstrom's (Jansson, 1984).

In spite of the optimality of this filter, it does have some disadvantages. First, it allows for negative contributions to the input estimation. Because negative processing load is supposed not to exist, this unrealistic estimation must be corrected. The filter also produces incorrect peaks. Without these incorrect peaks, the output estimation (using Equation A1) looks more like the measured output.

These disadvantages can be removed partially by two corrections. If $x_{est}(t)$ is the inverse Fourier transformed function of $X_{est}(f)$, $x_{est}(t)$ can be modified in two steps. First, the new estimation of the input is made to have no negative contributions:

$$x_{est,new,1}(t) = \max[0, x_{est}(t)]. \tag{A6}$$

In order to filter unrealistic peaks away, input contributions smaller than a particular fraction of the maximum are set to zero:

$$x_{est,new,fr}(t) = x_{est,new,1}(t) \quad x_{est,new,1}(t) > \text{fr} \cdot x_{max}$$

$$x_{est,new,fr}(t) = 0 \quad x_{est,new,1}(t) \leq \text{fr} \cdot x_{max} \tag{A7}$$

where fr is the fraction (i.e., 0.0, 0.1, 0.2, . . . , 0.9) and x_{max} is the maximal input value.

With each $x_{est,new,fr}(t)$, an estimate of the output can be calculated (using Equation A1). The $x_{est,new,fr}(t)$, which gives the best output estimate (according to a Euclidean measure), is taken for the ultimate input estimate.

The input estimate with the filter of Bracewell and Helstrom (Jansson, 1984) has one further disadvantage. The peaks of the input estimate are not very sharp. Because we assume that the input consists of peaked pulses only, we have developed a second method that uses both the filter of Bracewell and Helstrom and the least squares method to estimate the peaked input. We assume that the output has the following form:

$$y(t_n) = \sum_k c_k h(t_n - t_k), \qquad (A8)$$

where c_k is the input at time t_k (no delay is assumed).

With the least squares method, it is possible to calculate c_k. (Negative values of c_k are set to zero.) Then, t_k is calculated from the input estimation with the filter of Bracewell and Hel-strom, using the relative maxima of its input estimation. To calculate these maxima, some parameters have to be introduced:

$$x_{sum}(t_n) = \sum_{k \in \{1,n\}} x_{est}(t_k), \qquad (A9)$$

and the value of x_{sum} at the maximal time: $x_{sum,max} = x_{sum}(t_{max})$.

The position of the relative maxima are calculated according to the following procedure:

1. An input estimation is made with the filter of Bracewell and Helstrom.

2. The times $t_{0.01}$ and $t_{0.99}$ are determined, where $t_{0.01}$ is the biggest time with $x_{sum}(t) < 0.01 x_{sum,max}$ and $t_{0.99}$ the smallest time with $x_{sum}(t)/0.99 x_{sum,max}$. For $t < t_{0.01}$ and $t > t_{0.99}$, $x(t)$ is set to zero, and for other values of t, $x(t) = x_{est}(t)$.

3. The position of the relative maxima of $x(t)$ are calculated.