Parameter recovery for the time-window-of-integration (TWIN) model of multisensory integration in focused attention

Farid I. Kandil
School of Humanities and Social Sciences, Jacobs University Bremen, Bremen, Germany

Adele Diederich
School of Humanities and Social Sciences, Jacobs University Bremen, Bremen, Germany

Hans Colonius
Department of Psychology and Cluster of Excellence “Hearing4all,” Carl von Ossietzky University, Oldenburg, Germany

In multisensory settings such as the focused attention paradigm (FAP), subjects are instructed to respond to stimuli of the target modality only, yet reaction times tend to be shorter if an unattended stimulus is presented within a certain spatiotemporal vicinity of the target. The time window of integration (TWIN) model predicts successfully these observed cross-modal reaction time effects. It proposes that all the initially unimodal information must arrive at a point of integration within a certain time window in order to be integrated and thus to initiate response enhancements like the observed reaction time reductions. Here we conducted a parameter recovery study of the TWIN model for focused attention tasks, with five parameters (the durations of the visual and auditory unimodal and the integrated second stage, the width of the time window, and the effect size). Results show that parameter estimates were highly accurate (unbiased, constant error less than 5 ms) and precise (variable error less than 8 ms) throughout, speaking to a high reliability and criterion validity of the process. Further analyses ensured that the estimation procedure is consistent and sufficiently robust against contamination (faulty integration). It can thus be used to estimate reliably the point of integration and the width of the time window.

Introduction

Evidence for multisensory integration is found in many different formats, notably as either facilitation or inhibition of responses to crossmodal stimulus sets compared to unimodal stimulation. Prime examples of facilitation include an increased firing of a multisensory neuron in superior colliculus (Wallace, Wilkinson, & Stein, 1996), acceleration of manual or saccadic reaction time (Frens, Van Opstal, & Van der Willigen, 1995), and improved performance in audiovisual speech perception (Sumby & Pollack, 1954). The behavioral context considered here is the focused attention paradigm (FAP) where a visual and an acoustic stimulus are presented to the participants under specified spatiotemporal conditions (Figure 1). Their task is to only respond to the onset of a stimulus from a predefined target modality, such as the visual, for example, and to ignore the possible occurrence of a nontarget stimulus from the other modality, the auditory. It has nevertheless been shown that the nontarget stimulus can modulate the response, manual or saccadic, to the target. Depending on the spatiotemporal configuration of target and nontarget, this crossmodal effect results either in a speedup (facilitation) or a slowdown (inhibition) of the response (Amlôt, Walker, & Spence, 2003; Diederich & Colonius, 2007a). A prominent concept to describe the underlying dynamics is the spatiotemporal “window of integration” (e.g., Meredith, 2002). We have developed a quantitative model framework, the time-window-of-integration model (TWIN), to describe and predict reaction time (RT) modulation under crossmodal stimulation (Colonius & Diederich, 2004; Diederich & Colonius, 2012). A series of experimental studies found support for the main assumptions of the model (Diederich & Colonius, 2008a, 2008b; Colonius, Die-
The purpose of this study is to probe how well our model fitting algorithm is able to identify the numerical parameters controlling the model predictions. To this end, first, data are generated from the model using a range of model parameter values and, second, the parameters estimated from the simulated data are compared with the parameter values actually used for the simulated data (parameter recovery).

The TWIN model postulates that a cross-modal stimulus triggers a race mechanism in the very early, peripheral sensory pathways, which is then followed by a compound stage of converging sub-processes that comprise neural integration of the input and preparation of a response. Note that this second stage is defined by default: it includes all subsequent, possibly temporally overlapping, processes that are not part of the peripheral processes in the first stage. The central assumption of the model concerns the temporal configuration needed for multisensory integration to occur: multisensory integration occurs only if the peripheral processes of the first stage all terminate within a given temporal interval, the time window of integration (TWIN) assumption. Thus, the window acts as a filter determining whether afferent information delivered from different sensory organs is registered close enough in time to trigger multisensory integration. Passing the filter is necessary, but not sufficient, for cross-modal interaction to occur because the amount of interaction may also depend on many other aspects of the stimulus set, such as spatial configuration of the stimuli. The amount of cross-modal interaction manifests itself as an increase or decrease of second-stage processing time. Although this amount does not directly depend on the stimulus onset asynchrony (SOA) of the stimuli, temporal tuning of the interaction occurs because the probability of interaction is modulated by the SOA value. The five parameters to be recovered are (a) and (b) parameters \( \frac{k}{C_0} \) and \( \frac{A}{C_0} \) determining peripheral visual and auditory processing time, respectively; (c) parameter \( l \) representing the average second stage (central) processing time; (d) parameter \( x \) giving the width of the time window of integration; and (e) parameter \( D \) denoting the size of the cross-modal interaction effect in milliseconds.

The model framework encompasses both the FAP paradigm and the redundant signals paradigm (RSP) where, in the latter, stimuli from either modality are defined as targets (e.g., Colonius & Diederich, 2012). Here we focus on TWIN modeling of FAP only. For
this paradigm, it is additionally assumed that cross-modal integration only occurs if (a) a nontarget stimulus wins the race in the first stage opening the time window of integration, such that (b) the termination of the target peripheral process falls into the window. One interpretation is that a winning nontarget will keep the system in a state of heightened reactivity such that the upcoming target stimulus, if it falls into the time window, will trigger cross-modal interaction.1 If a stimulus from the target modality is the winner of the race in the peripheral channels, second-stage processing is initiated without any multisensory integration mechanism being involved.

In order to derive quantitative predictions for an empirical test of the model, further assumptions about the probability distributions of the various process durations postulated in TWIN have been introduced.2 These distributions are specified by numerical parameters that have to be estimated from the observed RT data. The model has been probed in a series of experiments under a variety of empirical conditions (Colonius et al., 2009; Diederich & Colonius, 2007b; Diederich & Colonius, 2008a). While these tests have generally supported the essential features of the model, the quality of a model test critically depends on one’s ability to obtain good estimates of the model parameters. The purpose of the current simulation study therefore is to probe whether, and how well, our parameter estimation method is able to recover the correct parameters. To this end, RT data sets are randomly generated from the model’s probability distributions with fixed and a-priori known parameter values. Next, estimates of these parameters are computed on these simulated data sets in order to check how well the original parameter values can be recovered. Obviously, the quality of the recovery results will depend both on properties of the model and on the efficiency of our estimation procedure. The next section provides details of the time window model and its simulation.

Time window of integration model

For concreteness, we assume the visual modality to be target modality and the auditory to be nontarget modality. The race in the first stage is represented by two statistically independent, exponentially distributed random variables \( V \) and \( A \) denoting the peripheral processing times for the visual and the acoustic stimulus with density

\[
f_\lambda(t) = \begin{cases} 
\lambda \exp[-\lambda t] & \text{if } t > 0 \\
0 & \text{if } t \leq 0,
\end{cases}
\]  

with \( \lambda > 0 \) and specific values \( \lambda \equiv \lambda_A \) and \( \lambda \equiv \lambda_V \) for the auditory and visual modality, respectively.

With \( \tau \) as SOA value and \( \omega \) as integration window width parameter, the above requirement for multisensory integration to take place is the realization of the event

\[ I_{\tau,\omega} = \{ A + \tau < V < A + \tau + \omega \}, \]

i.e., the nontarget stimulus wins the race in the first stage “opening the time window of integration” such that the termination of the target peripheral process \( V \) falls into the window. Here, a positive \( \tau \) value indicates that the visual stimulus is presented before the acoustic, and a negative \( \tau \) value indicates the reverse presentation order. Observable total reaction time is the sum of first stage processing time of the target modality, here \( V \), and second stage processing time, \( M \), assumed to follow a normal distribution and comprising all subsequent processes including motor preparation and execution and, in the bimodal condition, multisensory integration. The mean of the distribution of \( M \) differs depending on whether a unimodal or bimodal condition is considered. For unimodal trials, \( M \) has a mean of \( \mu \) and variance \( \sigma^2 \) (Figure 2A). For bimodal trials, second stage processing time depends on whether or not the condition for multisensory integration, \( I_{\tau,\omega} \), is met in a given trial (Figure 2B, C). Therefore, reaction time in the bimodal condition is mixture of two distributions with mean \( \mu \) and \( \mu - \Delta \), respectively, and mixing parameter \( P(I_{\tau,\omega}) \). Thus, mean reaction times in the unimodal and bimodal conditions are, respectively,

\[
E[RT_V] = \frac{1}{\lambda_V} + \mu
\]

\[
E[RT_{VA}] = \frac{1}{\lambda_A} + \mu - \Delta \times P(I_{\tau,\omega}), \tag{2}
\]

with \( \lambda_V \) denoting the exponential parameter for \( V \) introduced above and \( E \) the expectation operator (mean) of random variables. From Equation 2 it is obvious that when the cross-modal effect \( \Delta \) equals zero or when the probability of integration \( P(I_{\tau,\omega}) \) is zero, expected reaction time in the unimodal and bimodal condition will be identical.3 Probability of integration, \( P(I_{\tau,\omega}) \), is a function of the exponential parameters \( \lambda_V \) and \( \lambda_A \) for \( V \) and \( A \), the window width parameter \( \omega \), and the SOA value \( \tau \) that is determined by the experimental setup. Explicit expressions for \( P(I_{\tau,\omega}) \), depending on the sign of \( \tau \) and \( \tau + \omega \) can be found in the Appendix, see also Colonius and Diederich (2011).

Choosing the exponential distribution in TWIN is mainly motivated by computational simplicity. Adding a normally distributed second stage duration then results in the ex-Gaussian distribution for the observable reaction times, a distribution that has found some empirical support (Luce, 1986). Given the somewhat...
arbitrary choice of the ex-Gaussian, one would not expect the model to validly represent variability in the data and, thus, model fitting for TWIN has up to now been limited to the level of mean RTs.⁴

Parameter recovery and simulation of TWIN

Given an experiment with simultaneous presentation of visual target and auditory nontarget (i.e., $\tau = 0$) yields the following model equation for the bimodal means:

$$E[RT_{VA,0}] = \frac{1}{\lambda_V} + \mu - \Delta \times P(I_{0,0})$$

$$= \frac{1}{\lambda_V} + \mu - \Delta \times \frac{\lambda_A}{\lambda_A + \lambda_V} (1 - \exp[-\lambda_V\omega]).$$  (3)

There are five parameters to be estimated, $\lambda_A$, $\lambda_V$, $\mu$, $\omega$, and $\Delta$, but only two “observables” are available, i.e., the unimodal and bimodal sample means for $RT_V$ and $RT_{VA}$. Thus, the model and its parameters are clearly not identifiable. However, the situation can easily be remedied by increasing the number of SOAs. With four SOA values there will be model equations for four bimodal means and one unimodal mean, i.e., an equal number of “observables” and parameters (i.e., a “saturated” model). Importantly, however, note that here we are not investigating formal model identifiability of TWIN using theoretical approaches as developed, for example, in Bamber and van Santen (2000). Rather, we want to ascertain in a pragmatic way whether our parameter estimation method is able to recover the model parameter values from data sets that have been generated under the model with those same parameters. Before we present the details of the simulation procedure in the next section, two more issues have to be addressed.

First, we need to specify how to measure the “goodness” of the parameter recovery. In statistics, a
common measure for the quality of a finite-sample point estimator $\hat{\theta}$ for a parameter $\theta$ is the mean-squared error (MSE) defined by $E[(\hat{\theta} - \theta)^2]$. It is well known (Casella & Berger, 2002, p. 330) that then

$$E[(\hat{\theta} - \theta)^2] = Var[\hat{\theta}] + (E[\hat{\theta}] - \theta)^2, \quad (4)$$

where $(E[\hat{\theta}] - \theta)^2$ is (Bias $\hat{\theta}$)$^2$. Thus, MSE incorporates two components, $Var[\hat{\theta}]$ measuring the variability of the estimator (precision) and the other measuring its bias (accuracy) (Casella & Berger, 2002, p. 330). In analogy to this definition, we evaluated the precision and accuracy of the parameter recovery resulting from the simulations by considering the variability and central tendency of the parameter estimates across different simulated data sets. Moreover, we tested for the robustness of the parameter estimation method, that is, its ability to recover the original parameter values when participants committed certain errors like integrating although event $t_{int}$ was not realized (or vice versa).

The second issue concerns the generation of random variables in TWIN. Variability in the first stage is determined by the exponential parameters $\lambda_A$ and $\lambda_V$ that will be estimated by model fitting at the level of means. However, the variance of the normally distributed processing duration in the second stage, $\sigma^2$, is not included in this model fitting procedure and, therefore, cannot be “recovered.” On the other hand, in order to generate random data sets from the model, this parameter needs to be specified as well. We chose to fix the value at $\sigma = 25$ ms, a value that seems plausible given some experimental results obtained earlier (Diederich, Colonius, & Schomburg, 2008). Moreover, some pilot simulations suggested that the outcome of our main simulation study did not depend on the exact value of $\sigma$.

The main part of this evaluation study assessed accuracy and precision of the parameter estimation process in seven steps depicted in Figure 3: (a) For each of the five parameters of the basic TWIN model, we defined a set of parameter values. (b) We systematically combined them into 576 parameter vectors and (c) conducted a virtual experiment with 40 virtual subjects, all of them having that parameter vector as intrinsic variables. The experiment comprised Steps 4–6, in which reaction times were generated for each subject (subject level). Estimates across all subjects in an experiment were evaluated in comparison to the original parameter values for each experiment individually (experiment level). Finally, these evaluations were gathered across the 576 experiments to assess how well the parameter estimation process works if confronted with all sorts of parameter vectors, plausible, odd, and impossible ones (see below).

In additional experiments, we tested the consistency of the parameter estimation process, assessed the correlation between the parameters, and determined the degree of robustness of the estimation against contamination of the integration process. Finally, we investigated whether replacing the mean by the median in order to average the generated data (Step 4) would make the recovery more accurate or precise.

### Material and methods

The seven steps (Figure 3) will now be described in greater depth.

#### Step 1: Determining the parameter values

For each of the five parameters, three to four parameter values were used for the evaluation process (Table 1). They were selected to cover the entire range of (neurophysiologically) plausible values.

For both $\lambda^{-1}_{A}$ and $\lambda^{-1}_{V}$, i.e., the parameters for the lengths of the auditory and visual sensory stages, we assumed 20, 50, 100, and 150 ms to cover the range of possible values. If one assumes that the integration takes place in the cortex and that also the responses are triggered there (Molholm et al., 2002), then integration should not occur before the primary auditory and visual cortices (in macaque monkeys) are activated, i.e., some 20 ms after the auditory onset of noise (Camalier, D’Angelo, Sterbing-D’Angelo, de la Mothe, & Hackett, 2012) and some 30 ms after the visual onset (Schmolesky et al., 1998). On the other hand, responses to visual stimuli reach the frontal eye field, i.e., the secondary motor cortex for eye movements, between 60 and 100 ms after stimulus onset. We thus assumed twice that time as a safe candidate for the highest value.

The parameter values for $\mu$, the duration of the secondary (i.e., postintegration or common) stage, were simply obtained by subtracting the $1/\hat{\lambda}$ values mentioned above from a typical simple-reaction time of 200 to 250 ms: 50, 100, 150, and 200 ms. The parameter $\Delta$ stands for the temporal gain, i.e., the reaction time reduction, resulting from the integration of the multimodal information. Here, we took 20 and 50 ms to be reasonable values and added 100 ms as a safe maximum gain. Note that, for simplicity, we do not consider the case of inhibition, i.e., negative values of $\Delta$.

Parameter $\omega$ represents the width of the time window, i.e., the time interval during which integration may occur. In neurophysiological terms, this may be the duration for which the membrane potential of the integration cell remains excited. Using modern whole-
cell in-vivo recording methods, duration estimates vary greatly between very short, e.g., 2 ms for cells in the avian cochlea (Reyes, Rubel, & Spain, 1994), and long durations, e.g., 50–140 ms in hippocampus (Andersen, Raastad, & Storm, 1990). In contrast, area MT (V5) cells integrate motion information across more than 200 ms (Bair & Movshon, 2004). To cover the whole range, here we utilized values of 100, 200, and 300 ms.

**Step 2: Setting up the parameter vectors**

The parameter values defined in the first step were systematically combined into $4 \times 4 \times 4 \times 3 \times 3 = 576$ different vectors. Among them were plausible vectors such as (20, 50, 150, 200, 50), a vector that was identified in an earlier study for reaction time data using the same FAP paradigm (Diederich et al., 2008), as well as rather odd ones, which would represent gross imbalances between the values and the model parts they represent e.g., (20, 100, 50, 300, 150), and finally clearly *unrealistic* vectors such as (20, 20, 50, 200, 100), which would on average result in negative RTs if integration occurred (mean $RT = \lambda^{-1} + \mu - \Delta = 20 + 50 - 100 = -30$). All of these vectors were tested and taken into the evaluation, irrespective of their plausibility.
would be presented only with an acoustic stimulus and would have to react to it.

For each of the trials, pseudorandom numbers were generated for the first and second stage based on the distributions and the specific parameter values for the respective parameter. As described in the Introduction, \( A \) and \( V \) are random variables with exponential distributions (cf. Equation 1), while \( M \) is a random variable with a normal distribution \( N(\mu, \sigma = 25) \). As discussed above, the parameter \( \sigma \) was not recovered during the process and is therefore not part of this evaluation study.

The random realizations \( A, V, \) and \( M \) were then combined into an observable reaction time following the logic of the model:

\[
RT = \begin{cases} 
V + M - \Delta, & \text{if } \tau + A \leq V \leq \tau + A + \omega, \\
V + M, & \text{if } V < \tau + A \text{ or } \tau + A + \omega < V 
\end{cases}
\]

with \( \tau \) being the SOA value in the given condition. Reaction times were averaged across all the 200 trials in each condition.

**Step 5: Estimating the parameter values**

Mean reaction times for each of the nine conditions were handed over to the parameter estimation procedure that was programmed to find the best fitting estimates for the five parameter values. The objective was to minimize the objective function (Equation 6), i.e., the \( z \)-standardized differences between generated and predicted reaction times across all uni- and bimodal conditions,

\[
\text{Objective Function} = \sum_{\text{all conditions}} \left( \frac{\text{mean}[RT_{\text{gen}}] - RT_{\text{pred}}}{\text{standard error}[RT_{\text{gen}}]} \right)^2.
\]

The range of possible fits was thereby limited by the individual boundaries for the five parameters listed on the top line for each parameter in Table 1. The process started at random positions individually set anywhere between the lower and the upper boundary (for \( \mu \) between 5 and 300 ms).

**Step 6: Evaluating the results on the experiment level**

In all experiments and in all subjects, the average difference between generated and predicted mean RTs was about 6 ms, indicating that the process
could find some vector that explained the data sufficiently well (cf. Figure 4).

**Step 7: Assessing accuracy and precision on the study level**

These estimate differences were gathered across the 40 subjects and 576 experiments and evaluated in terms of accuracy and precision. This evaluation was performed separately for each parameter (cf. Figure 5), and within each parameter both across all parameter values (leftmost insets) and for each parameter value individually (right insets). Accuracy refers here to the absolute distance between the original parameter value and the average parameter estimate. A high accuracy is indicated by a low distance and thus a small constant error, and indicates in turn a high criterion validity. Precision, on the other hand, measures the (smallness of the) width of the estimates’ distribution. A high precision comes with a small variable error (typically the standard deviation or variance), and indicates in turn a high reliability. A high reliability is the prerequisite for a high validity, i.e., accuracy can only be evaluated if the standard error is sufficiently low.

All simulations and model fittings were ran in Matlab using the Matlab function fminsearch (The MathWorks, Inc., Natick, MA, USA). All statistics were performed using Mathematica (Wolfram Research, Inc., Champaign, IL, USA).

**Results**

**Recovery results on subject, experiment, and study level**

As mentioned above, results in this study are described on three levels, the subject level (\(n = 1\) subject), the experiment level (\(n = 40\) subjects forming a group), and the study level (across all 576 experiments and thus \(n = 23,040\) subjects). While the main questions of this evaluation concern the study level, good and reliable results on subject and experiment levels are the necessary precondition for the analysis of the main results presented further below.

Results on the subject level are shown in Figure 4 (Panels A and B). Mean RTs generated for each subject were used to estimate the parameters for the subject. Panel A shows the means of the generated reaction times for one person for the seven bimodal and the two unimodal conditions (gray symbols, \(n = 200\) trials per condition) along with the RTs predicted by the model based on the estimated parameters. The differences between these reaction times were fairly small, with approximately 1.9 ms per condition for the plausible parameter vector. As depicted in Panel B, across all \(n = 40\) subjects in all \(n = 576\) experiments, this prediction error had a mean of 5.95 ms (SD = 3.64) per condition.

Panel C shows the results on the experiment level. Evaluation on this level comprised analyzing the results of all subjects participating in a single experiment (please note that, for the sake of clarity, this figure shows a result from a pretest, in which 500 subjects were used rather than the usual 40). The panel shows the estimate statistics of the five parameters for one sample experiment (plausible parameters). About one half of the estimates met the original parameter value with an error of less than about 15 ms (box in the boxplots), whereas the other half of the estimates deviated from that original value by more than \(\approx 15\) and up to some 50 ms (antennas in the boxplots). Only a minor fraction (1%–2%) between experiments deviate by more than 50 ms.

However, outliers are not the only possible issue. The recovery could also yield invalid estimations if the process terminated in or near the limiting boundaries. This can best be seen in the leftmost boxplot in Figure 4C, in which the lower antenna ends exactly at 5 ms, i.e., the lower boundary. This means that a significant proportion of the estimates in these experiments is invalid. Cumulative results across the recoveries for all subjects in all experiments are given in the bar plots of Figure 4D and in Table 1. Black bars indicate the percentage of valid estimates for the given parameter and parameter value, while the gray bars to its left or right, where existing, give the percentage of the recoveries that terminated in the lower or upper boundary, respectively. (Boundaries are indicated in that figure by the small black stick-like bars at the ends of the scale.) Results show that while more than 90% of the estimates are valid, some 10% are invalid, with a higher occurrence of the latter for very low parameter values (for \(\lambda_A, \lambda_Y, \mu,\) and \(\Delta\)), which are by themselves very close to the boundary and thus endangered from the outset, or very high values (for \(\omega\)).

For the study level, individual parameter estimates were averaged using the median across the 40 subjects of an experiment into group estimates. All of these group estimates stood clear off the boundaries and thus all further assessments (and the graphs shown in Figure 5) rely on the full amount of 576 estimates. Results for the evaluation of the recovery are shown in Figure 5 and Table 2. Graphs in Panels A to E show histograms for the five parameters, and while the leftmost graphs show histograms aggregated across all parameter values, the others (to the right) show them for each parameter value individually. To ease comparison, histograms are plotted against the distance of the recovered parameter to the original value.
For parameters $\lambda_A^{-1}$, $\lambda_V^{-1}$, and $\mu$, the histograms ubiquitously are rather narrowly distributed around the expected value (zero on the x axis). Medians (constant errors) of the group estimates for the first three parameters ($\lambda_A^{-1}$, $\lambda_V^{-1}$, $\mu$) range between $-2.5$ ms and $+2.5$ ms (means are between $-5$ and $5$ ms) for individual parameter values, and level out across all parameter values to about half that size, thereby exhibiting a very high accuracy of the estimates. Median absolute deviations (MAD) are lower than $7$ ms ($SD = 16$ ms), speaking to a high precision. Given the low absolute distances and the three times higher deviation measures, the estimates have to be classified as unbiased.

For parameter $\omega$, the same assessment (narrow distribution, unbiasedness) holds true for the first two parameter values of 100 and 200 ms. Although the medians are negative, they are close to zero (medians $> -5$ ms; means $> -10$ ms), while the median absolute deviation is again lower than 7 ms ($SD < 24$ ms). In contrast, for $\omega = 300$ ms, the constant error is higher (median $= -13$ ms, mean $= -18$ ms) and the distribution is much wider (MAD $= 35$ ms, $SD = 82$ ms) and shallower than in all other cases.

Estimates for parameter $\Delta$ (Figure 5, Panel E) show, on first glance, a stable overestimation (85% of the estimates are positive), which could speak in favor of a biasedness.
However, the constant (median) error is less than 5 ms (mean < 8 ms), and the variable error is comparable to the ones found for the first three parameters (or even less) with median and standard deviations of approximately 3–4 and 8–10 ms, respectively.

Further evaluation results

Consistency

As sample size per experimental condition increases, one would expect that the parameter estimate based on the sample will improve. In statistical theory, an estimator is called consistent if it converges to the true (population) value when sample size becomes infinitely large. Note that a consistent estimator may be biased and an unbiased estimator may not be consistent. For many estimators, like maximum likelihood estimators, consistency holds under rather general conditions. However, the parameter estimation method utilized here does not guarantee consistency of the estimators. On the other hand, consistency follows if both the bias of the estimator and its variance converge to zero for sample size approaching infinity (Lehmann, 1983). Therefore, we tested the accuracy and precision behavior of our estimates with increasing sample size.
We used the standard plausible parameter vector \( (k/C_0, A, k/C_0 V, l, x, D) = (20, 50, 100, 200, 50) \), with the standard number of 40 subjects and with a number of trials per condition (sample size) increasing from 40 to 200 in steps of 20 and from 200 to 1000 in steps of 100 trials. Results in Figure 6 show that the bias for the median decreases with the square root of the sample size (black symbols) and that same holds true for the median absolute deviation (gray symbols) for all five parameters. This strongly suggest that the parameter estimation procedure produces consistent estimators.

### Correlation between parameter estimates

A potential problem in recovering parameters for a model is that the value of the estimate for one parameter may be significantly correlated with the value of another (Ratcliff & Tuerlinckx, 2002). In trying to optimize the fit, the fitting procedure may trade off the value of one parameter against the value of another one. In the TWIN model, one would expect this to happen to some degree between the \( \lambda \) parameters of the first stage and the mean \( \mu \) of the second stage processing. This issue becomes especially critical when one tries to interpret differences among parameter values across experimental conditions. For example, increasing stimulus intensity should be reflected in differences in the \( \lambda \) parameters but not in \( D \) because intensity affects first stage processing and integration probability, but not the amount of facilitation or inhibition in the second stage, according to TWIN assumptions.

Parameter correlation was tested again using the standard plausible parameter vector \( (20, 50, 150, 200, 50) \), 40 virtual subjects and an increasing number of trials (sample size) per condition. The expectation was that both the constant error (mean and median) and the standard error (standard and median deviation) decrease with the sample size. Results for both error measures show such a pattern suggesting consistency of the estimators. Median differences are shown in black and median absolute deviation in gray.

### Table 2. Main results of the study. Accuracy and precision of the recovered parameters are given for each parameter separately: first across all parameter values (All), then for each parameter value individually. Accuracy (unbiasedness) refers to a preferably low constant error, i.e., difference from the original parameter (mean or median), whereas a high precision is indicated by a low variable error (standard deviation, SD, or median absolute deviation [MAD]).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Mean Diffs</th>
<th>SD</th>
<th>Median Diffs</th>
<th>MAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 1/\lambda_A )</td>
<td>All</td>
<td>-0.25</td>
<td>4.50</td>
<td>-1.94</td>
<td>12.10</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.75</td>
<td>2.85</td>
<td>1.85</td>
<td>6.87</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>-0.52</td>
<td>3.40</td>
<td>-1.19</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>-2.25</td>
<td>5.65</td>
<td>-4.81</td>
<td>13.35</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>-0.75</td>
<td>5.85</td>
<td>-3.59</td>
<td>15.07</td>
</tr>
<tr>
<td>( 1/\lambda_V )</td>
<td>All</td>
<td>-0.73</td>
<td>4.03</td>
<td>-2.77</td>
<td>12.61</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>-1.03</td>
<td>2.03</td>
<td>-1.34</td>
<td>4.33</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>-1.95</td>
<td>4.20</td>
<td>-4.21</td>
<td>10.27</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>1.00</td>
<td>5.43</td>
<td>-1.02</td>
<td>15.04</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.00</td>
<td>6.45</td>
<td>-4.55</td>
<td>16.70</td>
</tr>
<tr>
<td>( \mu )</td>
<td>All</td>
<td>1.35</td>
<td>3.58</td>
<td>3.20</td>
<td>11.97</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>1.20</td>
<td>3.68</td>
<td>3.49</td>
<td>14.26</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>1.48</td>
<td>3.30</td>
<td>3.18</td>
<td>12.13</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>1.05</td>
<td>3.78</td>
<td>2.79</td>
<td>9.69</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>1.70</td>
<td>3.93</td>
<td>3.33</td>
<td>11.49</td>
</tr>
<tr>
<td>( \omega )</td>
<td>All</td>
<td>-2.82</td>
<td>8.05</td>
<td>2.26</td>
<td>51.54</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>-1.50</td>
<td>2.95</td>
<td>-4.32</td>
<td>9.08</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>-4.05</td>
<td>5.65</td>
<td>-9.71</td>
<td>23.59</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>-6.75</td>
<td>34.68</td>
<td>20.73</td>
<td>82.48</td>
</tr>
<tr>
<td>( \Delta )</td>
<td>All</td>
<td>3.65</td>
<td>3.60</td>
<td>6.71</td>
<td>9.25</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>4.70</td>
<td>3.20</td>
<td>7.62</td>
<td>8.07</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>2.78</td>
<td>3.18</td>
<td>6.41</td>
<td>10.05</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>2.75</td>
<td>3.65</td>
<td>6.09</td>
<td>9.48</td>
</tr>
</tbody>
</table>
parameters along. There is a very high correlation of $r > 0.98$ between $\lambda_A^{-1}$ and $\lambda_V^{-1}$ and a comparably high but negative correlation between $\lambda_A^{-1}$ and $\mu$ and between $\lambda_V^{-1}$ and $\mu$; further a strong, positive correlation between the estimates for the first two parameters and $\Delta$ with $r > 0.88$, and a comparably high but negative correlation between the estimates for $\mu$ and $\Delta$, with $r = -0.87$; finally a moderate ($r = 0.38$) but significant negative correlation between $\omega$ and $\Delta$. Correlations between $\omega$ and the first three parameters were not significant. While some of these correlations reflect expected trade-off effects of the estimation procedure, they also point to some caveats for interpreting parameter differences between experimental conditions.

**Robustness against contamination**

Contaminated reaction times come from some process other than the one being studied. For example,
the subject may occasionally anticipate the stimulus rather than processing it or there may be some distraction from the task by various sources not under the control of the experimenter. A number of general methods, e.g., the truncation of data sets, have been considered to deal with such spurious RTs (Ulrich & Miller, 1994). Here we limit discussion to the case where, in a certain percentage of trials, subjects violate the integration mechanism postulated in TWIN by exhibiting the RT reduction of Δ milliseconds when the condition for integration was not satisfied, or vice versa. Robustness of the parameter estimation procedure, that is the accuracy and precision of the parameter recovery, was tested by systematically increasing the proportion of such faulty integration behavior. In general, i.e., in the unbiased case (i, see below), data were generated so that integration and the associated reduction in reaction time took place in only (100 – p)% of the due cases, whereas RT was not reduced in the other p%. At the same time, reaction times were generated incorporating the reduction in p% of the cases for which integration should not have taken place. In six experiments, the percentage p was systematically increased from 0% to 25% in steps of 5%, using the same plausible parameter vector (20, 50, 150, 200, 50) as before. Generated RTs were averaged and presented to the parameter estimation procedure.

We assessed the robustness of the parameter estimation in respect to four questions: How robust is the process against (a) unbiased and (b) biased contamination of the integration? (c) Is the robustness additionally affected by the width of the temporal window (ω)? And (d) can the robustness of the parameter estimation be increased by using the median over the mean in averaging the RT data?

Results are shown in the panels in Figure 8. Black lines represent the constant error (accuracy) in terms of the difference between the median across the parameter estimates across the 40 subjects, whereas gray lines represent the variable error (precision), i.e., the median absolute deviation. Generally, parameter estimates become less accurate and precise with an increasing percentage of contaminated trials. We defined a difference of 20 ms from the original parameter value as a threshold for all parameters (i.e., one standard deviation in the unbiased case in Figure 8A at 0% contamination) in order to compare the multiple conditions tested here among each other.

Data for question (a) show (central panel in Figure 8A) that the process is generally very finely tuned and highly sensitive to even a low contamination of the RT data. Approximately 10%–15% sufficed for the mean group estimates to reach the threshold. The pattern of estimation error, i.e., whether the parameter is over- or underestimated is clear and consistent even for small amounts of contamination: \( \lambda_A^{-1}, \lambda_V^{-1} \), and Δ become overestimated, whereas μ and ω become underestimated. However, while accuracy worsens with contamination, the variable error (gray lines), and thus the precision, remains stable across a wide range of contamination, meaning that the error in parameter estimation is quite systematic.

With respect to question (b), results show that contamination can have opposing effects, depending on the kind of bias the subjects have (left and right panel in Figure 8A). If only integration trials are affected by the contamination, then the subject exerts a bias towards nonintegration. As a consequence, contamination does not seem to have any early effect on the estimates (left panel). Only if contamination is higher than 15%, Δ becomes underestimated, which, however, is a consequence from the decreasing average RT gain. In contrast, if only the nonintegration part is contaminated, subjects have a bias towards integration. In this case the errors seen in the central panel increase by factor of two and consequently the thresholds are reduced by factor of two for all parameters.

In question (c) we investigated whether the ω value had any effect on the robustness, since we had observed some gross under- and overestimates for ω in the main analyses. We therefore assessed robustness for three values of ω. While we fixed \( \lambda_A^{-1}, \lambda_V^{-1}, \mu, \) and Δ to the values used before (standard vector), we increased ω from 100 to 200 (standard value) to 300 ms. Results (Figure 8B) for the first three parameters seem to be rather constant across the three panels and thresholds do not differ. The only effect the increase seems to have, is a systematic decrease of the threshold for ω associated with an earlier underestimation of ω. The effect resembles the pattern observed for the ω estimates in Figure 5D. However, while this is true for the deviations in absolute measures, the relative deviation from the original parameter value becomes lower and lower with the increasing ω. These effects are paralleled by a general overestimation of Δ, which is additionally reduced with increasing ω.

Question (d) investigated the possibility to make the parameter estimation more robust against contamination. Because subjects tend to be inattentive to the task or anticipate the response in a certain proportion of trials, this contamination, and especially biased contamination, constitutes a real rather than just a theoretical problem. One way of diminishing the impact of these spurious (too long or too short) reaction times is to use the median rather than the mean in order to average the 200 single RT data into the mean/median reaction times before these are handed over to the parameter estimation procedure (Step 4 in the overview shown in Figure 3). These modifications diminish the constant error by factor four and increases the threshold by factor two, and thus
Figure 8. Robustness of the parameter recovery against contamination. For the plausible parameter vector (20, 50, 150, 200, 50), graphs show dependence of the parameter estimates on the proportion $p$ of contaminated response generations, i.e., when subjects do or do not integrate in trials in which they wouldn’t or would have to integrate. Graphs show the constant error (median difference
make the process much more robust against contamination.

**Discussion**

**Recovery results**

This study evaluated the parameter estimation method in the time-window-of-integration (TWIN) model of multisensory integration (Colonius & Diederich, 2004) for its ability to recover the underlying true parameter values. To this end, different values for five model parameters \((\lambda_A^{-1}, \lambda_V^{-1}, \mu, \omega, \Delta)\) were systematically combined into a total of 576 different parameter sets (vectors of length five) for each of which a virtual experiment was conducted involving 40 subjects producing 200 responses per experimental condition. Data were generated by sampling from the probability distributions postulated in the model, best-fitting parameter values were obtained by minimizing an objective function defined for the model, and the resulting distributions of parameter estimates were evaluated with respect to accuracy (constant error) and precision (variable error).

Our data reveal that the model parameters are recovered with very high accuracy, i.e., with low constant errors (median differences to the original value) of less than 5 ms, and also very high precision, i.e., low variable errors (median absolute deviations) of less than 7–8 ms for practically all parameters and their values. It should be stressed that the range of parameter vectors was huge, comprising neurophysiologically plausible vectors, such as \((20, 50, 100, 200, 50)\), as well as odd ones like \((20, 100, 50, 300, 150)\), which would represent gross imbalances among the values and thus among the model parts they represent; and even unrealistic (forbidden) vectors, such as \((20, 20, 50, 200, 100)\), which would, if integration occurred, lead to negative reaction times.

The only exception from these unambiguously good results has been observed for the highest value of parameter \(\omega\). Further investigation revealed that the distribution of estimates for \(\omega = 300\) ms can be divided into two subpopulations (Figure 9A). For \(\lambda_A^{-1} \leq \lambda_V^{-1}\), the distribution of \(\omega\) estimates is rather narrow (MAD = 11 ms). In contrast, for the complementary case, \(\lambda_A^{-1} > \lambda_V^{-1}\) (black bars), the estimates have a wide distribution (MAD = 43 ms).

This odd but systematic pattern can be understood by inspecting the dependence of integration probability \(\mathcal{P}(I)\) on the intensity parameters \(\lambda_A\) and \(\lambda_V\) across the SOA range (cf. the equations in the Appendix). The corresponding graphs in Figure 9B depict the two opposite cases relating to the two subpopulations defined by \(\lambda_A^{-1} < \lambda_V^{-1}\) (left-hand graph) and \(\lambda_A^{-1} \geq \lambda_V^{-1}\) (right-hand graph). Each graph presents a family of ordered probability of integration curves corresponding to seven different values of \(\omega\), from \(\omega = 100\) ms (lightest gray) to \(\omega = 700\) ms (darkest gray). Obviously, for \(\omega\) values of 200 ms and above the curves on the right (where \(\lambda_A^{-1} \geq \lambda_V^{-1}\)) collapse for most SOA values, whereas for the other subpopulation the probability of integration curves still differ for different values of \(\omega\).

As a consequence, because the objective function controlling the estimation procedure directly depends on the probability of integration, actually, the term \(\mathcal{P}(I) \times \Delta\) the relative insensitivity of \(\mathcal{P}(I)\) with respect to the exact \(\omega\) value for higher window widths makes determination of optimal \(\omega\) values less precise for the subpopulation defined by \(\lambda_A^{-1} \geq \lambda_V^{-1}\).

**Consistency, correlation, and robustness of estimates**

In additional experiments, we investigated consistency, correlation between recovered estimates, and robustness/sensitivity of recovery against contamination, and we compared using the mean or the median for data averaging at the different levels of the parameter estimation process.

**Consistency**

The parameter estimates provided by the estimation process become more accurate and precise with an increasing number of trials per condition, and thus with an increasingly more reliable average reaction time for each of the conditions. Both the constant error (difference of the mean from the original parameter) and the variable error (standard deviation) decrease
The high negative correlation between the first-stage intensity parameters $\lambda_V$ and $\lambda_A$ and second stage parameter $\mu$ is not surprising given the additivity of the two stages at the mean RT level. A similar argument holds for the high negative correlation between $\omega$ and the (positive) $\Delta$ values and the moderate negative correlation between $\omega$ and $\Delta$. These relationships must be kept in mind when interpreting parameter values across different experimental conditions. On the other hand, the absence of correlation between window width $\omega$ and parameters $\lambda_V$, $\lambda_A$, and $\mu$ is encouraging for testing certain hypotheses. For example, one may probe the inverse-effectiveness hypothesis by comparing experimental conditions with high and low intensity values. If a larger window width is associated with low intensity $\lambda$ parameters, and vice versa, a conclusion in favor of that hypothesis may be less prone to an artifact of correlations among the parameter estimates.

Robustness

Robustness of the parameter estimation procedure was tested assuming that, in a certain percentage of trials, the rule for integration was violated, i.e., that stimuli not registered within the time window were nevertheless integrated (perhaps due to a strong temporal offset of the stimuli) or that subjects failed to integrate acoustic and visual information that should have been integrated.

Parameter estimation turned out to be robust if the subject does not integrate faulty in more than 10%–15% of the trials (cf. Figure 8A, central panel). However, this holds true only for unbiased subjects who behave faulty in both ways, i.e., integrate when they are not to, and vice versa, with the same probability. In contrast, if they have a bias towards nonintegration, contamination does not seem to have any effect on the good quality of the parameter estimation (left panel), whereas the effects of contamination are significantly increased for subjects with a bias towards integration. For these subjects, the threshold goes down to 5%–9% (right panel).

Importantly, sensitivity to contamination can be met effectively by a minor change in the data pooling process, that is, before the parameter estimation begins. If one averages the single reaction times of a subject using the median rather than the mean, and replacing the standard deviation (SD) by the median absolute deviation (MAD) in the objective function, contamination does not contribute to such a great extent (Figure 8C). Thus for future research, we recommend to use the objective function (Equation 6) in this modified version.

Estimating the point of integration in the processing stream?

Despite the constraints in regard to the $\omega$, the parameter estimates can be used to calculate the point of integration in the (cortical) processing stream and the gain with a confidence interval of $\pm 33$ ms; and if only one experiment is run, the width of the time window of integration (TWIN) can be estimated with a C.I. of $\pm 70$ ms. Still, as the time from first activation of a cell in V1 (latency of 45 ms after stimulus onset) to that of the last cell in FEF (latency of 100 ms) is only 55 ms (Schmolesky et al., 1998), our parameter estimates cannot be used to decide upon the localization within the visuo-saccadic processing stream.
Even if one used the estimates from all 576 experiments to probe localization, with the high accuracy (median of 2.5 ms) and the high precision (median deviation of 7 ms) as discussed before, good localization would not be possible. From Figure 10, the median firing latencies of cells in V1, V3, V5 (MT), and FEF, which all lie along the presumed cortical processing stream of a visually induced saccade, differ by a mere 10 ms, while their distributions have a wide standard deviation of another 10 ms, and thus overlap to a great extent. Thus, even for an ideal point estimate with 0 ms constant error and 0 ms variable error, no reliable distinction could be made between these stages in the processing stream. However, estimations obtained here may help decide whether multisensory integration takes place directly in the brain stem, thalamus, occipital/parietal cortex, or rather late motor cortex.

**Conclusion**

The time window of integration (TWIN) model, as developed over the last 10 years by Colonius, Diederich, and colleagues constitutes, on the one hand, a general framework describing how multisensory integration unfolds over time; on the other hand, by adding specific statistical assumptions about the distributions of the various postulated sub-processes, it allows us to predict quantitative changes of multisensory integration effects in reaction time, as a function of the specific experimental conditions like spatiotemporal stimulus configuration, instruction of the participants, or personal variables like age. While the general framework does not involve the issue of parameter estimation, the specific distributional instantiations of the model will only be useful for hypothesis testing if accuracy and precision of the parameter estimates can be assured. Importantly, a satisfactory fit of a model to the data alone does not guarantee that the parameters can be interpreted meaningfully. Let us consider a situation where one wishes to determine the actual durations of the primary auditory/visual stage and of the second, central stage. For example, assume estimates for the first stage parameters ($\lambda_A^{-1}$ and $\lambda_V^{-1}$) are systematically overestimated by 50 ms, while average second stage processing time ($\mu$) is systematically underestimated by the same amount. As a consequence, one might falsely localize the point of integration away from early processing areas towards motor areas in the brain. While we did not set a fixed threshold, given a total reaction time in experiments with real subjects of 200 to 250 ms, we considered errors of up to 10 ms as acceptable. Precision of the parameter estimates may even be of more importance. For example, assume one investigates whether the time window is larger (wider) in one condition compared to another condition or if the window width measured for a subject differs from the optimal window, with optimality derived from a Bayesian argument (Colonius & Diederich, 2010; Diederich et al., 2008). If the window estimates ($\omega$) for the two conditions are 100 and 150 ms, say, one needs to know about the precision (the reciprocal of the variance) to infer whether this difference is real or just a product of chance.

To conclude, we have demonstrated that the parameter estimation procedure recovers the parameters in the TWIN model with a very high accuracy (unbiasedness) and precision, speaking to a high reliability and criterion validity of the process. The goodness-of-fit procedure provided consistent parameter estimates, without problematic dependencies (correlations) among the parameter estimates. Furthermore, using either the mean or the median in order to average (a) single reaction time data in the objective function, or (b) to average parameter estimates for individual subjects into a (grand) group estimate, constitutes an adjustment wheel to make the parameter estimation process either sensitive, balanced, or robust against faulty integration (contamination) in subjects, depending on what is required in a given experimental study.

**Keywords:** parameter recovery, multisensory processing, crossmodal integration, focused attention paradigm, time window of integration, TWIN model
Acknowledgments

We thank Achim Gelessus for help with using Jacobs University’s computer cluster (CLAMV). Work on this study was supported by the Deutsche Forschungsgemeinschaft (DFG) grant DI 506-12 (AD and HC) and grant SFB-TRR-31 “Active Listening” (HC).

Commercial relationships: none.
Corresponding author: Farid I. Kandil.
Email: f.kandil@jacobs-university.de.
Address: School of Humanities and Social Sciences, Jacobs University Bremen, Bremen, Germany.

Footnotes

1For saccadic eye movements, this may correspond to a gradual inhibition of fixation neurons (in superior colliculus) and/or omnipause neurons (in midline pontine brainstem).
2Some distribution-free model tests have been performed as well; see Diederich and Colonius (2008b).
3Note that the parameter of the auditory peripheral processing time only enters this equation via $P(I_{1,ao})$. One might argue that termination of first stage processing should be determined by the minimum of $V$ and $A$. This is a plausible model version but experimental tests did not find strong support for it (Colonius & Arndt, 2001).
4However, Colonius and Diederich (2013) developed a first analysis of RT variability in TWIN.
5Including implausible and unrealistic parameter values presents an additional stress test of the model. While a high degree of flexibility of a model to accommodate even implausible parameter combinations may be considered a drawback, it should be realized that the model does not know about their implausibility and, as long as plausible parameters are recoverable, there is no reason to be worried.

References


Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.


### Appendix

**Probability of integration as function of SOA and window width**

The peripheral processing times $V$ for the visual and $A$ for the acoustic stimulus have an exponential distribution with parameters $\lambda_V$ and $\lambda_A$, respectively. Thus,

$$f_V(t) = \lambda_V e^{-\lambda_V t},$$

$$f_A(t) = \lambda_A e^{-\lambda_A t},$$

for $t \geq 0$, and $f_V(t) = f_A(t) = 0$ for $t < 0$. The corresponding distribution functions are referred to as $F_V(t)$ and $F_A(t)$, respectively. The visual stimulus is the target and the acoustic stimulus is the nontarget. By the model

$$P(I_{t,\omega}) = Pr(A + \tau < V < A + \tau + \omega) = \int_0^\infty f_A(x)\{F_V(x + \tau + \omega) - F_V(x + \tau)\}dx,$$

where $\tau$ denotes the SOA value and $\omega$ is the width of the integration window. Computing the integral expression requires that we distinguish between three cases for the sign of $\tau + \omega$:

(i) $\tau + \omega < 0$

$$P(I_{t,\omega}) = \int_{-\tau-\omega}^{-\tau} \lambda_A e^{-\lambda_A x} \{1 - e^{-\lambda_V (x + \tau + \omega)}\}dx$$

$$+ \int_{-\tau-\omega}^{-\tau} \lambda_A e^{-\lambda_A x} \{e^{-\lambda_V (x + \tau)} - e^{-\lambda_V (x + \tau + \omega)}\}dx$$

$$= \frac{\lambda_V}{\lambda_V + \lambda_A} e^{\lambda_A \tau} (-1 + e^{\lambda_V \omega});$$
(ii) $\tau < 0 < \tau + \omega$

$$P(I_{t,\omega}) = \int_0^\tau \lambda_A e^{-\lambda_A x} \left\{ 1 - e^{-\lambda_V(x+\tau+\omega)} \right\} \, dx$$

$$+ \int_\tau^\infty \lambda_A e^{-\lambda_A x} \left\{ e^{-\lambda_V(x+\tau)} - e^{-\lambda_V(x+\tau+\omega)} \right\} \, dx$$

$$= \lambda_A \left\{ \frac{1}{\lambda_V + \lambda_A} \left( 1 - e^{-\lambda_V(x+\tau)} \right) + \lambda_V \left( 1 - e^{\lambda_V \tau} \right) \right\}.$$

(iii) $0 < \tau < \tau + \omega$

$$P(I_{t,\omega}) = \int_0^\infty \lambda_A e^{-\lambda_A x} \left\{ e^{-\lambda_V(x+\tau)} - e^{-\lambda_V(x+\tau+\omega)} \right\} \, dx$$

$$= \lambda_A \left\{ \frac{\lambda_V}{\lambda_V + \lambda_A} \left( e^{-\lambda_V \tau} - e^{-\lambda_V (\omega+\tau)} \right) \right\}.$$