

Cross-modal integration of simple auditory and visual events

GEOFFREY R. PATCHING

University of Nottingham, Nottingham, England

and

PHILIP T. QUINLAN

University of York, York, England

Responses are typically faster and more accurate when both auditory and visual modalities are stimulated than when only one is. This bimodal advantage is generally attributed to a speeding of responding on bimodal trials, relative to unimodal trials. It remains possible that this effect might be due to a performance decrement on unimodal ones. To investigate this, two levels of auditory and visual signal intensities were combined in a double-factorial paradigm. Responses to the onset of the imperative signal were measured under go/no-go conditions. Mean reaction times to the four types of bimodal stimuli exhibited a superadditive interaction. This is evidence for the parallel self-terminating processing of the two signal components. Violations of the race model inequality also occurred, and measures of processing capacity showed that efficiency was greater on the bimodal than on the unimodal trials. These data are discussed in terms of a possible underlying neural substrate.

In the attempt to understand how the different components of multimodal stimuli are processed, a basic paradigm has been developed in which simple reaction times (SRTs) are measured to the onset of various types of imperative signals. Such an imperative signal can be a single visual stimulus, a single auditory stimulus, or the concurrent presentation of both of these stimuli (see, e.g., Hershenson, 1962; Miller, 1982). For expository convenience, a *unimodal stimulus* is defined as one in which either an auditory or a visual signal is presented alone, and *bimodal stimulus* is defined as one in which both auditory and visual signals are presented concurrently. The typical finding is that RTs to the bimodal stimuli are shorter than the RTs to either type of unimodal signal; This basic effect will be termed the *bimodal advantage*. For example, Miller (1982) instructed participants to respond whenever a simple tone or light was presented but to withhold responding when neither was presented. He reported a seemingly typ-

ical pattern of results—namely, RTs to the bimodal stimulus were shorter (326 msec) than those to either the single auditory signal (409 msec) or to the single visual signal (412 msec).

Empirically, there is no doubt about the existence of such bisensory effects. However, there has been controversy over whether such effects should be taken as evidence for independence of processing of auditory and visual events or whether such effects reveal interesting properties about how the different sensory-processing streams interact in their operations. A primary concern involves mathematical models, which can account for the bimodal advantage without positing any interaction between auditory and visual processes (Raab, 1962; Townsend & Nozawa, 1997).

In an early paper, Raab (1962) argued that a bimodal advantage can arise without the need to posit any interaction between auditory and visual processing. He demonstrated that, provided participants respond to the first detected component of the bimodal stimulus and detection time is variable, a bimodal advantage will occur on the basis of statistical sampling alone. On this view, auditory and visual processes proceed quite independently of each other, and responses are based on whichever of these processes is completed first (Raab, 1962). This account has become known as *statistical facilitation*. Models based on statistical facilitation alone are often referred to as *independent race* models, because auditory and visual processes independently race to evoke a response (see Mordkoff & Yanits, 1991, for a review). As Miller (1982) has pointed out, such models make the strong prediction that the RT on bimodal trials cannot be shorter than the shortest RT on uni-

The work reported here was, in part, supported by a Research Studentship funded by the Biotechnical and Biological Sciences Research Council, U.K., to the first author. James Townsend and J. Toby Mordkoff provided extremely helpful reviews on an earlier version of this article. Correspondence regarding this work may be addressed to either G. R. Patching, Department of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, England (e-mail: grp@psychology.nottingham.ac.uk) or P. T. Quinlan, Department of Psychology, University of York, Heslington, York YO10 5DD, England (e-mail: ptq1@york.ac.uk).

Note—This article was accepted by the previous editorial team, headed by Neil Macmillan.

modal trials. He went on to propose a particularly powerful test by which to assess the plausibility of such models, now commonly referred to as *race model inequality*. Race model inequality adopts a stringent criterion by which to compare RT on bimodal trials with that predicted on the basis of statistical facilitation alone.

On this analysis, it is assumed that the different signal components are dealt with by separate information-processing channels. A detection response can be elicited when processing is completed on either channel. As a consequence, RTs may be described by distributions of random variables x_1 and x_2 , where x_i is the time needed for processing to be completed on channel C_i . By the independent race account, the RT is simply the minimum of x_1 and x_2 . This places a strict constraint on responding on redundant signal trials. More formally, this constraint is defined as follows:

$$P(\text{RT} < t | C_1 \text{ and } C_2) \leq P(\text{RT} < t | C_1) + P(\text{RT} < t | C_2).$$

The term on the left of the inequality refers to the cumulative probability of making a response on a redundant signal trial by the given time t . The first term on the right of the inequality refers to the cumulative probability of making a response by t when Signal Component 1 is presented (alone), and the second term refers to the cumulative probability of making a response by t when Signal Component 2 is presented (alone).

What the inequality sets out is that, according to any simple race account of processing on the separate channels, there is a strict lower bound to RTs on the redundant signal trials that is fixed by the speed of processing on the single-component signal trials. For instance, RTs to a bimodal signal cannot be shorter than the shortest RT recorded on either of the unimodal signal trials. More important, violation of the race model inequality rules out any account of performance that is based on the notion of simple independent racing (Miller, 1982; Mordkoff & Yantis, 1991). Indeed, studies of the bimodal advantage have consistently claimed to provide good evidence that falsifies independent process models based on statistical facilitation alone (see, e.g., Diederich & Colonius, 1987; Gielen, Schmidt, & Van den Heuvel, 1983; Giray & Ulrich, 1993; Hughes, Reuter-Lorenz, Nozawa, & Fendrich, 1994; Miller, 1982, 1986).

To account for any violation of race model inequality, Miller (1982) argued in favor of a class of models termed *coactivation* models. Such models rest on the assumption that activations from both signals are combined and that this combined activation builds gradually over time. When such summed activation reaches a criterion level, a response can be executed. On this account, there is an important sense in which information from across the auditory and the visual modalities is integrated; in other words, the different processing streams are seen to interact in their operations.

Although the popularity of models based on coactivation has grown since Miller (1982) first set out the race model inequality (Mordkoff & Yantis, 1991; Townsend &

Nozawa, 1995), the present article was motivated by a desire to test possible alternative reasons for the basic bimodal advantage. For example, it seemed necessary to attempt to rule out a more mundane reason for the bimodal advantage than that provided by coactivation models—namely, that the bimodal advantage arises because of the slowing of RT on unimodal trials, relative to bimodal ones. This is of some import, since researchers have occasionally expressed concern that the bimodal advantage might arise not so much as a result of a speeding of RT on bimodal trials but, rather, because of a slowing of RT on unimodal trials (see, e.g., Bernstein, 1970; Fournier & Eriksen, 1990; Nickerson, 1970, 1973; Townsend & Nozawa, 1997).

An additional problem also arises in considering the possibility of modality-switching costs that may be present on unimodal trials (Quinlan & Hill, 1999; Spence & Driver, 1997). Such costs are defined as the performance decrement that occurs across trials when the imperative stimulus changes modality. In the standard SRT paradigm, discussed above, both unimodal and bimodal stimuli are intermixed. Regardless of which type of stimulus has been presented on the trial immediately prior to the presentation of a bimodal stimulus, there will be no cost due to modality switching. In contrast, if two unimodal stimuli are presented in immediate succession in different modalities, modality-switching costs may occur. Hence, the difference between performance on bimodal and unimodal trials may be due, in critical respects, to a slowing on the unimodal trials. In this regard, it is interesting to note that the bimodal advantage has been found to be smaller in studies in which the different unimodal and bimodal signals have been presented in different blocks of trials, relative to that when the same signals were intermixed within the same block of trials (see Giray & Ulrich, 1993). In such blocking paradigms, performance on unimodal trials cannot be contaminated by modality-switching costs, simply because the different unimodal stimuli are presented in separate blocks of trials. On unimodal trials, participants may focus their attention purely on the critical modality.

Other grounds for concern over the nature of the bimodal advantage stem from the arguments recently put forward by Townsend and Nozawa (1997). They have shown, on the basis of mathematical modeling, that a variety of independent exhaustive models can give rise to a bimodal advantage of sufficient magnitude to violate race model inequality. Townsend and Nozawa (1997) advocated a class of models in which signal components are processed one at a time and, furthermore, in which both processing *channels* (or *modalities*, in the present context) are always checked (hence, the claim of exhaustive processing; see also Grice, Canham, & Boroughs, 1984; Grice, Canham, & Gwynne, 1984). To rule out independent exhaustive models, Townsend and Nozawa (1997) advocated the use of what they referred to as the *double-factorial paradigm*.¹ Essentially, this paradigm is an outgrowth of the additive factors method of Sternberg (1969). This analysis is based on patterns of interaction and additivity in designs in which

the joint effects of two two-leveled factors are investigated. The results are evaluated by analysis of variance (ANOVA), where additivity is thought to provide strong evidence of serial processing (see Nozawa, Reuter-Lorenz, & Hughes, 1994). Townsend and Nozawa (1997) discussed the case in which Factor 1 relates to presence versus absence of a signal in one modality and Factor 2 relates to the speed of processing of the separate signals. However, in an earlier paper (Townsend & Nozawa, 1995), they were clear that any two factors may be appropriate as long as they can be seen to affect separate processes.

In order to adapt the double-factorial methodology for use with the basic SRT paradigm discussed here, the two factors used were the stimulus intensities of the auditory and the visual signals. Component stimulus intensity has been shown to be effective in altering performance in the SRT task. For example, Hughes et al. (1994) used a simple manual response task in which the intensities of the auditory and visual signals were varied and RTs to the different-intensity auditory and visual signals were collected within the same block of trials. Essentially, high- and low-intensity auditory signals (i.e., bursts of white noise) were defined, together with high- and low-intensity visual signals (i.e., illumination of LEDs). Generally speaking, performance showed effects of signal intensity, with shorter responses being made when high-intensity signals were presented than when low-intensity signals were presented. Unfortunately, though, Hughes et al. did not provide enough information to ascertain what the relationship was between signal intensity and RT on bimodal trials. All they stated was that “[t]here was little effect of mismatches between the auditory and visual detection times on the size of the redundant-targets effect” (Hughes et al., 1994, p. 141).

The present study reports on an experiment in which the component stimulus intensities of the auditory and the visual signals were factorially combined within a block of trials that contained both unimodal and bimodal stimuli. The aim of the experiment was to attempt to provide evidence that might help adjudicate between the different accounts of the bimodal advantage that have just been discussed.

Of fundamental concern is what happens on bimodal trials following factorial combination of component signal intensity. Three possible outcomes of these four cell means have been discussed in the literature (Nozawa et al., 1994; Townsend & Nozawa, 1995).² The first possible outcome is one in which a simple additive effect is found between the two levels of auditory and visual intensity: Statistically significant main effects of the separate signal intensities are discussed in the absence of a statistically reliable interaction. Such an outcome would be evidence in favor of a serial and independent stage model of processing. That is, the processing of one of the signals is completed prior to the processing of the other signal.

In contrast to such an additive effect of the two factors, it is possible to describe two further cases in which interactions arise. The first, covered in some detail by Nozawa et al. (1994), is the case of overadditivity, or superadditiv-

ity. With this case, statistically significant main effects of signal intensity are coupled with an interaction that shows that participants are particularly impeded in responding when the low-intensity values of the signals are combined. A superadditive result is most consistent with the notion of some form of self-terminating, parallel processing of the two components on bimodal trials. Nozawa et al. argued that such a pattern of results is consistent with the idea that participants are responding to the first signal to be processed and, as such, fits with accounts of the bimodal advantage that are based on either coactivation or statistical facilitation.

The final outcome discussed is one in which statistically significant main effects of signal intensity are coupled with an interaction that shows subadditivity, or underadditivity. As Townsend and Nozawa (1995, p. 325) have pointed out, such a pattern of underadditivity fits most comfortably with a system that embodies exhaustive parallel processing. In the present context, such an outcome would be evidence that participants process the auditory and the visual components concurrently but that a response cannot be emitted until processing has run to completion on both modalities. At first glance, such a state of affairs appears to be a quite unlikely account of the bimodal advantage. For instance, if the independent-processing account is accepted, participants should be able to respond as soon as they detect a signal in either modality. Alternatively, if the coactivation account is accepted, then as soon as the summed signal energy reaches some response criterion, the participant should be able to respond. These accounts of performance are, however, contradicted by the predictions of the exhaustive parallel model, as described clearly by Egeth and Dagenbach (1991).³

By their framework, it is possible to discuss the three following cases: (1) the case in which the two high-intensity components are combined (i.e., the high/high case), (2) the case in which the two low-intensity components are combined (i.e., the low/low case), and finally, (3) the case in which a low-intensity component is combined with a high-intensity component (i.e., the high/low case). In their example, the two components were separate visually presented letters, and the intensities of the two letters were treated as being the factor under discussion. By the logic described by Townsend and Schweickert (1989) and Schweickert and Townsend (1989), Egeth and Dagenbach (1991) stated that if the mean time to complete the processing of a high-intensity signal is T msec and the mean time to complete the processing of a low-intensity signal is $T + DT$ msec, it follows that predicted RT on the high/high trials is T msec, on the low/low trials is $T + DT$ msec, and on the high/low trials is also $T + DT$ msec. Such predictions are exact if the corresponding intensities of the associated signals are equated. Other cases can be discussed in which the corresponding high and low values of the different signals are not equated, but nevertheless, the basic account predicts that a response on any given trial cannot be made until processing has been completed for the slower signal. What is clear, though, is that an under-

additive interaction appears to rule out both coactivation and independent race accounts of performance on the bimodal trials.⁴

Overall, therefore, the double-factorial design appears to provide a useful test of the various competing models of the bimodal advantage. On these grounds, this particular paradigm was used in an attempt to provide differential support for one of the accounts.

METHOD

Participants

Twenty participants from the York Departmental Participant Panel took part in the experiment. The participants received course credits or payment for taking part. All reported normal vision, and none reported any hearing problems.

Apparatus

The experiment was controlled by an Apple Macintosh 7200 series computer. The visual stimuli were presented on a 25.5×20 cm video monitor; the auditory stimuli were presented via two Sony speakers (Model SRS-A91) placed either side of the video monitor, 30 cm from the center of the screen. Stimulus presentation and timing were controlled with a PsyScope script (Cohen, MacWhinney, Flatt, & Provost, 1993). RTs were collected via the PsyScope buttonbox. The experiment was conducted in a quiet darkened room. In addition, a chinrest was fixed at a distance of 50 cm from the screen of the monitor, to ensure a constant viewing distance.

Stimuli and Design

The participants were required to make a simple detection response to the presentation of a unimodal auditory signal, a unimodal visual signal, or a bimodal signal. The visual signal was a centrally presented asterisk on the screen of the monitor. The auditory signal was a tone presented simultaneously through both speakers. Two sorts of asterisks were used: a bright one (Courier 48 point) and a dim one (Courier 14 point). Viewed from a distance of 50 cm, the bright asterisk subtended a visual angle of 0.92° and measured 0.8 cm in extent; the dim asterisk subtended a visual angle of 0.23° and measured 0.2 cm in extent. To ensure that an effect of visual intensity obtained on the trials with a single visual signal, the dim asterisk was made dimmer by increasing the amount of gray in its hue. The screen background was black, and the asterisks were colored white. The intensities of the bright and the dim asterisks were 0.9 and 0.02 cd/m², respectively. Auditory targets were sinusoids, synthesized digitally with a sampling rate of 44.1 kHz. Two sorts of tones were used: a loud one presented at 70 dB SPL and a soft one presented at 40 dB SPL. The frequency of both tones was 640 Hz. Each trial was initiated with the central presentation of a Courier 14-point exclamation mark on the screen of the monitor.

Both the loud and the soft tones and both the bright and the dim visual asterisks and their factorial combinations were randomly intermixed within the same block of trials. The participants were required to complete three of these blocks of trials in one experimental session. Each block of trials comprised 20 loud tones, 20 soft tones, 20 bright asterisks, and 20 dim asterisks; factorially combined, these produced four different bimodal stimuli: loud/bright, loud/dim, soft/bright, and soft/dim, each of which was presented 20 times. In addition, 20 no-go trials were presented, in which no stimulus appeared following offset of the fixation mark.⁵ The order of all these trial types was randomized within each block.

Procedure

Prior to the presentation of each block of trials, the participants were presented with written instructions on the monitor, informing

them to respond as quickly and as accurately as possible to the presentation of a tone, an asterisk, or a tone and an asterisk but to withhold their response when no stimulus was presented. These instructions were read to them, and the importance of responding as quickly and as accurately as possible was stressed. Each trial began with the presentation of a visual warning signal (an exclamation mark) for 250 msec. If the trial was a blank trial, no other stimulus was presented. However, for all the other trials, 250 msec after the offset of the fixation mark, the asterisk, the tone, or both of these signals were presented for 150 msec. On signal trials, the participants had to make a response within a 1,000-msec response deadline. If a participant responded within 1,000 msec of the offset of the fixation mark on a blank trial, this was recorded as an error.

RESULTS

Summaries of the RT data are presented in Figure 1. In arriving at these values, all responses less than 100 msec were considered as anticipations and were discarded from the analysis of the RTs. In addition, the mean percentage of responses to blank trials (i.e., false alarms) was computed for each of the conditions. Misses were defined as cases in which a signal was presented but the participant failed to make a response within the 1,000-msec response window. The participants were highly accurate: 0.1% of the responses were classified as anticipations, 6.7% as false alarms, and 1.3% as misses. There was no evidence of any speed/accuracy tradeoffs in the data.

Analyses of the data will be described in two parts. In the first, overall mean performance for responses to the unimodal and bimodal stimuli will be considered. The manipulations of stimulus energy on the unimodal trials gave rise to independent effects for the separate visual and auditory signal components. This shows that the manipulations of the energy of the separate signal components was effective. More important, though, are the data for the bimodal signal trials, for here the mean RTs describe a superadditive interaction across the four trial types.

In the second part, more detailed analyses will be reported of performance for the four different types of bimodal stimuli. Here, the initial aim was to compare the data with the race inequality in a bid to see whether performance on the bimodal trials provided any differential support for coactivation accounts of performance. As will be shown, the data did provide differential support for such accounts, but the most compelling evidence arose in the data for the trials in which the signal energy was at a maximum. Finally, an alternative but complementary approach to data analysis will be described, in which the processing capacity on the bimodal trials is tested further. These analyses provide some additional support for efficiency gains on bimodal trials.

Mean Performance

Unimodal trials. The RT data from the unimodal trials were entered into a 2×2 repeated measures ANOVA. The two within-subjects factors were trial type (auditory vs. visual) and signal intensity (high vs. low). The analysis revealed a statistically significant main effect of trial

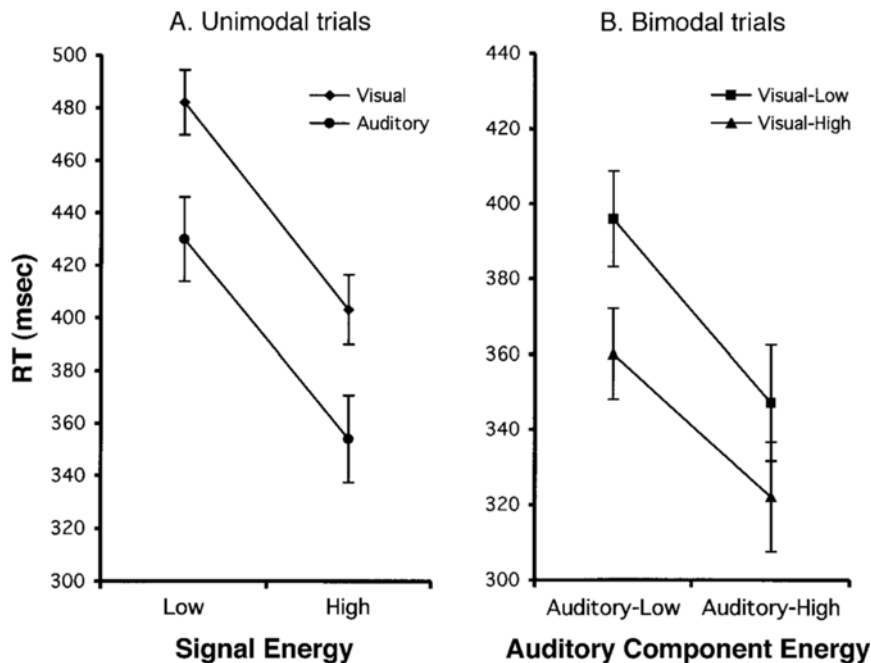


Figure 1. Mean reaction times (RTs) for different stimulus types for (A) the unimodal and (B) the bimodal stimulus trials. Error bars reflect standard errors of the corresponding means.

type [$F(1,19) = 43.01, p < .001$]: The participants responded more quickly to auditory than to visual signals (392 vs. 443 msec, respectively). In addition, the main effect of signal intensity was also found to be statistically significant [$F(1,19) = 160.53, p < .001$]: the participants responded more quickly to high- than to low-intensity signals (379 vs. 456 msec, respectively). The interaction of trial type and signal intensity was not found to be statistically reliable [$F(1,19) = 0.15, p > .05$].

Bimodal trials. The RT data for the bimodal trials were entered into a 2×2 repeated measures ANOVA. The two within-subjects factors were auditory component intensity (loud vs. soft) and visual component intensity (bright vs. dim). This analysis revealed statistically significant main effects of both auditory component intensity [$F(1,19) = 63.45, p < .001$] and visual component intensity [$F(1,19) = 102.67, p < .001$]. In addition, the interaction between these two factors was statistically reliable [$F(1,19) = 6.57, p < .05$].

To examine this pattern of performance in more detail, the following mean interaction contrast was computed for each participant:

$$\overline{RT}(D,S) - \overline{RT}(B,S) - \overline{RT}(D,L) + \overline{RT}(B,L).$$

These four means have been computed from the data for the trials in which the dim/soft, bright/soft, dim/loud, and bright/loud stimuli were presented, respectively. This interaction contrast has been discussed in detail by Townsend and Nozawa (1995) and provides critical evidence about the nature of the underlying interaction. Critically, a re-

sulting negative value indicates an underadditive interaction, whereas a resulting positive value indicates an overadditive interaction. This contrast was, therefore, computed for all the participants, and these data were analyzed via a one-sample (two-tailed) t test. This test revealed that, across participants, the contrast was statistically greater than zero [$t(19) = 2.56, p < .05$, mean value = 12.33]. In this respect, the data rule out any account that assumes either serial processing or exhaustive parallel processing of the stimulus information in the different modalities.

Testing the Data Against the Race Inequality

It is, nevertheless, important to try to adjudicate between a simple statistical facilitation account of the data and the alternative coactivation account. To do this, the data were analyzed and tested against the predictions of the race model inequality. Figure 2 shows the cumulative distribution functions (CDFs) of RTs for each of the four types of bimodal trials, together with the CDFs for the sum of the comparable RT data for the unimodal trials. Miller (1982) has argued that to be consistent with race models, the CDF obtained on bimodal trials should be everywhere below and to the right of the sum of the corresponding unimodal CDFs. Visual inspection of Figure 2 shows that violations of this rule appear to have occurred for the loud/bright trials through the 5th–35th percentiles (see Figure 2). After Miller (1982), paired t tests were conducted across participants at each of the percentile points for these functions. The results of these analyses showed RTs on loud/bright trials to be reliably shorter than the

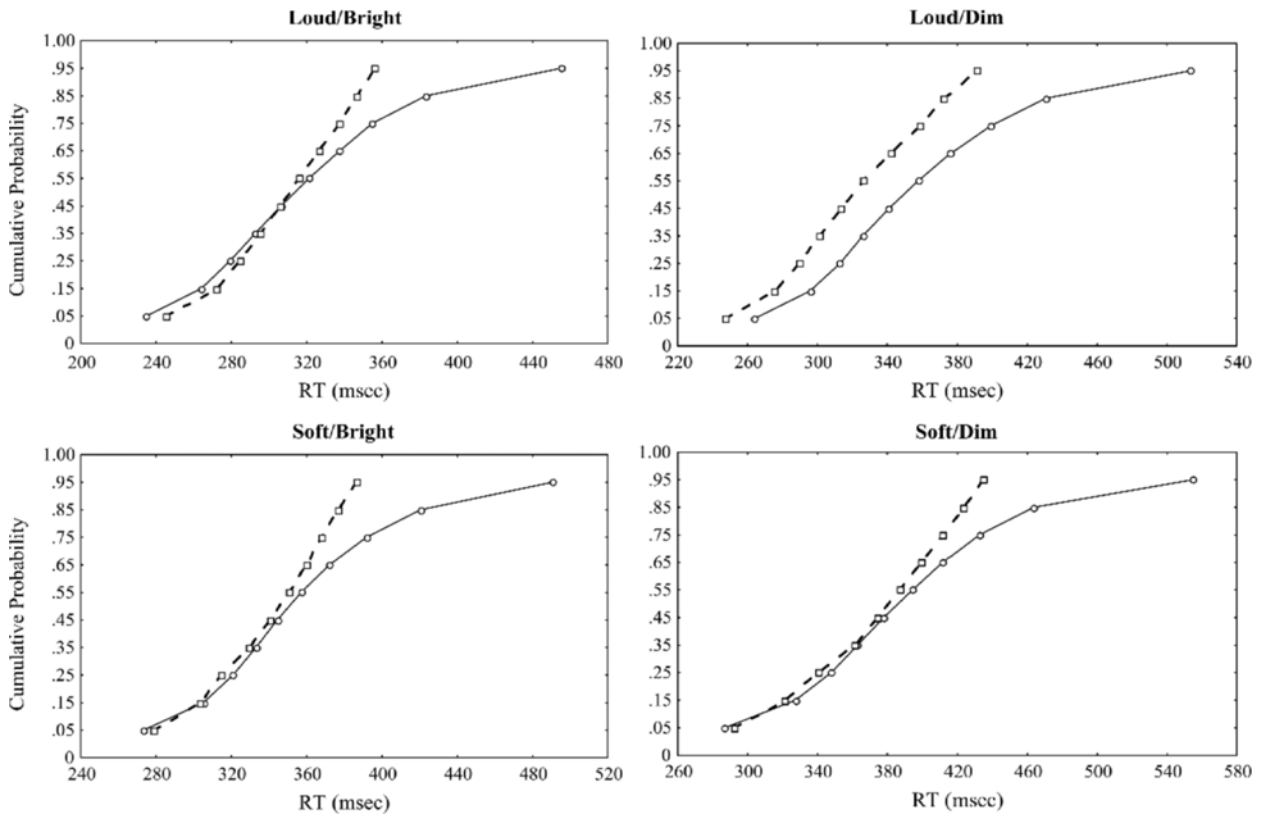


Figure 2. Graphs showing the cumulative distribution function (CDF) for the sum of the unimodal trials, together with the CDF for the corresponding bimodal trials for each type of bimodal stimulus. The dashed line denotes the summed distribution, and the solid line the bimodal distribution.

corresponding RTs from the sum of unimodal CDFs at the 15th and 25th percentile points [$t(19) = 1.80$ and $t(19) = 1.85$, $p < .05$, respectively, one-tailed tests]. This pattern of performance shows that the bimodal advantage cannot be adequately accounted for on the basis of statistical facilitation alone. The data from the loud/ bright trials are consistent with the postulate that information from across the auditory and visual modalities is pooled prior to a response: There is evidence of processes of coactivation.

Considerations of Processing Capacity

Various notions of processing capacity have been discussed in the literature, but perhaps the most explicit have been put forward by Townsend and colleagues. For instance, Townsend and Nozawa (1997) discussed notions of processing capacity in tasks similar to those described here, where the term *capacity* was used to relate variations in task load to efficiency of processing. Wenger and Townsend (2000) discussed the same notion in terms of how much energy a system expends in accomplishing any particular set of operations. At a more formal level, central to this approach is the concept of a *hazard function*. Such a function provides an estimate of the probability of finishing a task in the next instant of time t , given that the task

is not yet completed. Integrating this function results in the *integrated hazard function*, and this in turn provides a measure of the total amount of energy needed to complete a given task by some critical time t . It is possible, therefore, to derive integrated hazard functions for the different types of imperative signals used in the simple sorts of speeded tasks described here. What is of particular interest, in this context, is the definition of processing capacity as given by the *capacity coefficient* $C(t)$. $C(t)$ is derived by dividing the integrated hazard function for responses to a particular type of bimodal signal by the sum of the integrated hazard functions for responses to the component signals on the corresponding unimodal trials.⁶

Given this metric of processing capacity, Wenger and Townsend (2000) have discussed three possible outcomes. The first is one in which the sum of the component unimodal integrated hazard functions is greater than that produced for the corresponding bimodal stimulus. In this case, $C(t) < 1$, indicating that participants' capacity for processing each component signal is reduced when another must be processed at the same time. The second outcome is one in which the sum of the unimodal integrated hazard functions is equal to that for the bimodal stimulus. In this case, $C(t) = 1$, and this reflects the fact that each

component can be processed just as efficiently when presented as part of a bimodal stimulus as when presented alone. The final case is one in which the sum of unimodal integrated hazard functions is less than that for the bimodal stimulus. In this case, $C(t) > 1$, and this obtains when the signals in combination are processed more efficiently than when either is presented alone.

Following the procedures laid down by Wenger and Townsend (2000), values of $C(t)$ were computed for each participant in 10-msec RT bins for each of the four bimodal trial types. These data are shown in the four graphs

that constitute Figure 3. Examination of Figure 3 shows that for each of the four types of bimodal trials, the values of $C(t)$ exceeded 1 across the whole range of RTs. That is, there is clear evidence that the system exhibited supercapacity. Importantly, of the 20 participants tested, all but 1 produced values of $C(t) > 1$ for each of the four bimodal trial types. The distinctive participant produced values of $C(t) > 1$ for three of the four bimodal trial types; the loud/dim case was the exception. Overall, therefore, the data fit with the idea that the perceptual system worked more efficiently when the separate senses were stimulated to-

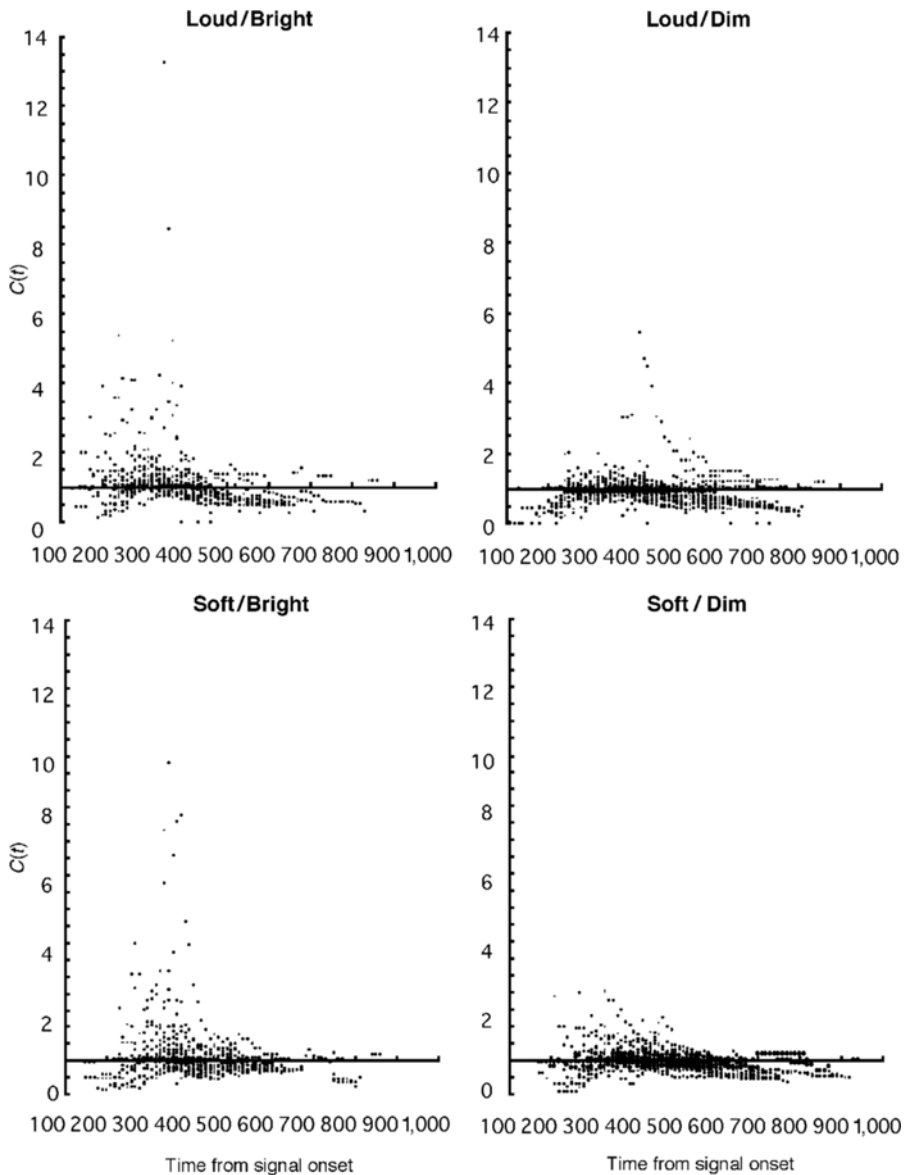


Figure 3. Values of the capacity coefficient $C(t)$, for each of the bimodal stimulus types computed in 10-msec bins across the complete range of reaction times from the onset of the stimulus across all participants.

gether than when either sense was stimulated alone. In this way, the data reveal that the participants' performance was in line with the notion of supercapacity as defined by Wenger and Townsend (2000).

GENERAL DISCUSSION

The primary motivation behind the present study was to attempt to rule out accounts of the bimodal advantage that contradict claims about how such an effect informs about the cross-modal integration of auditory and visual information. For instance, one such account (Townsend & Nozawa, 1997) assumes that participants slavishly check both modalities on every trial and, because of this, RTs are predicted to be longer on unimodal than on bimodal trials. Such an account is important because it attempts to explain the presence of a bimodal advantage in the absence of making any claims about cross-modal integration. Overall, however, the present data provide little support for such an account and are far more suggestive of some form of interaction in the processing of concurrently presented auditory and visual signals.

Three approaches to data analysis have been reported. In the first, overall mean performance with the various types of stimuli was scrutinized. The experiment was based on the factorial combination of high and low levels of stimulus energy for auditory and visual signals. The participants made simple go/no-go responses on each trial and were instructed to respond only when any signal occurred. The analysis of the unimodal signal trials revealed additive effects of both visual and auditory stimulus energy. The participants responded more quickly to a high-energy signal than to a low-energy signal; this held true for both auditory and visual modalities. Therefore, the data clearly showed that the underlying processes were sensitive to the amount of stimulus energy present on a given trial.

More important, though, was performance on the bimodal stimulus trials. The critical result here was the presence of a superadditive interaction across the means for the four combinations of signal energy. According to Townsend and colleagues (Nozawa et al., 1994; Schweickert & Townsend, 1989; Townsend & Ashby, 1983; Townsend & Nozawa, 1995; Townsend & Schweickert, 1989), such a result provides a strict constraint on theorizing about the bimodal advantage. The superadditive interaction in the present pattern of performance suggested that some form of self-terminating or coactive parallel processing was responsible. As Nozawa et al. (1994) have pointed out, this general pattern is consistent with both a race account of performance and a channel summation account.

In an attempt to adjudicate between these two accounts, further analyses were carried out to see whether any violations of Miller's (1982) race inequality were present in the data. Here, it was found that violations of the race model did occur, but only with the bimodal loud/bright signals. Such violations of race inequality are typically taken to suggest the operation of some form of coactivation, wherein information integration across the two modalities takes

place prior to a response. The superadditive interaction, together with the evidence of coactivation, fits with accounts of information processing in which the two modalities operate in parallel and some form of information integration takes place prior to a response. Although there are issues relating to the actual locus of where such information integration may be taking place (in the sequence of stages from sensory registration to response), the present data fail to throw any further light on these issues. (For a much more thorough discussion of these issues, see Mordkoff, Miller, & Roch, 1996.)

The present experimental data stand in contrast to the alternative modality-switching accounts that posit slowing of unimodal RTs, relative to bimodal RTs. The switching accounts are clear that responses on bimodal trials cannot exceed that predicted by statistical facilitation—that is, RTs based on the shorter of the completion times for the single modalities. Such a prediction is contradicted by the evidence suggesting coactivation in the data. This is not to argue that such switching costs are not present in the data but merely that such an account cannot provide an adequate explanation for the overall pattern of results. To categorically rule out channel-switching accounts, though, further data need to be collected. Although the present data suggest that "true" coactivation does take place, the results are also compatible with the type of interactive race processes discussed by Mordkoff and Yantis (1991). Closure on this issue can come about only when the contingencies for responding have been systematically varied in tandem with the current manipulations of signal intensity (see Mordkoff, 1998, for an example of how altering response contingencies can help adjudicate between opposing theoretical alternatives). Such a possibility remains an interesting avenue for future research.

The final approach to data analysis concerned the derivation of capacity coefficients for all the participants across the four bimodal stimulus types (after Townsend & Nozawa, 1995). Here, the critical finding was that all the participants exhibited supercapacity and all, except one, exhibited supercapacity for all four types of bimodal stimulus. The present data therefore revealed more efficient processing when both signal components were present on a trial than when either alone was present. Although supercapacity was revealed for all four bimodal stimulus types, evidence of coactivation arose only with one of these. Coactivation, as shown by the violations of the race inequality, occurred only for the loud/bright stimulus—that is, the stimulus containing the maximum amount of stimulus energy. Nevertheless, both sets of analyses converge to produce the conclusion that participants are particularly efficient when presented with a bimodal signal comprising high levels of component signal energy.

The failures to show violations of Miller's inequality in the presence of supercapacity are perfectly in line with the claim made by Townsend and Nozawa (1995) that a system might at some point exhibit supercapacity but also fail to violate Miller's inequality (p. 336). Indeed, as has been noted previously, on more than one occasion (Eriksen,

Goettl, St. James, & Fournier, 1989; Fournier & Eriksen, 1990; Mordkoff & Yantis, 1991), Miller's inequality may place too strict a condition on theorizing and may even be "overly severe" (Eriksen et al., 1989). For instance, as Mordkoff and Yantis (1991) have pointed out, failures to find violations of the inequality cannot in themselves rule out coactive systems. In this regard, the dissociations between the indices of processing capacity and the bimodal advantage may merely reflect the conservative nature of Miller's inequality.

Throughout, it has been assumed that the violations of the race inequality reflect the operation of mechanisms responsible for cross-modal integration. Moreover, although it has been assumed that such integrating mechanisms are operating whenever a bimodal stimulus is present, the operation of such mechanisms is, here, apparent only when the stimulus energy is considerable. At the neural level, evidence is growing for the existence of both modality-specific and cross-modal sensitive mechanisms. Whereas the notion of separate auditory and visual systems is beyond doubt, more recent research has pointed to the existence of cells in the superior colliculus that are taken to integrate information across the auditory and the visual modalities (Stein & Meredith, 1990, 1993; Stein, Meredith, Honeycutt, & McDade, 1989). Neural summation provides a plausible account of the bimodal advantage obtained in the present experiment, and the data indicate that such processes are critically influenced by the amount of stimulus energy present on a given trial.

Following previous work, Nozawa et al. (1994) arrived at similar conclusions on the basis of an application of the double-factorial paradigm in which eye movements, rather than manual responses, were measured. They, too, manipulated stimulus energy for auditory and visual signals and were therefore able to generate bimodal signals defined relative to the factorial combinations of high and low energy on these two dimensions. Participants were instructed to generate a saccade to the location of the imperative signal. On a given trial, the stimulus could occur at random at one of two contralateral locations. On catch trials, no stimulus was presented. There is much agreement across their results and those of the present study, but there are also some points of contrast. For instance, in their paper, although the superadditive interaction in mean performance for responses to the four types of bimodal stimuli failed to reach statistical significance, there was reliable evidence of coactivation for all four types of stimuli. In this regard, they argued strongly that some form of cross-modal integration taking place within the saccadic control system.

However, given that eye movements were irrelevant to the present task, there is no reason to assume that exactly the same mechanisms underlie performance across the different tasks. Indeed, coactivation in the present task arose only for the bimodal stimulus with the greatest energy. Nevertheless, there is reasonable agreement across the two studies for evidence of coactivation, and similar effects due to manipulations of stimulus energy have also been shown. Both sets of data suggest that processing

across the two modalities occurs in parallel and that task performance reveals evidence of cross-modal integration.

In closing, it is interesting to speculate about possible relations between the putative modality-specific and the cross-modal integrating mechanisms. Although there is some evidence to suggest a direct relationship between the magnitude of the bimodal advantage and component signal detection times, the present findings indicate that this is not necessarily so. For instance, in the present study, performance with the bimodal loud/bright and soft/bright stimuli can be seen to be equally efficient, but in the former condition RTs to comparable intensities single-signal trials are mismatched, and in the latter condition matched. As such, these data conform to those obtained by Hughes et al. (1994) in showing little effect of mismatches between unimodal auditory and visual RTs on the efficiency with which the bimodal signals could be processed. On these grounds, it seems that there is a degree of independence between the modality-specific and the cross-modal mechanisms. It may, therefore, be a mistake to assume that the integrating mechanisms simply sum the outputs from dedicated visual and auditory input channels. Indeed, when detailed analyses of multisensory neurons in the superior colliculus have been carried out (Stein & Meredith, 1993), it has been shown that the presentation of a bimodal stimulus can result in a response that is significantly greater than that produced when either unimodal stimulus is presented in isolation. Moreover, such response enhancement can exceed that predicted by summing the two corresponding unimodal responses.

Although any attempt to forge a link between neurophysiology and behavior at such a level may seem premature, the present data do accord rather well with what is currently known about the underlying neural substrate.

REFERENCES

- BERNSTEIN, I. H. (1970). Can we see and hear at the same time? Some recent studies of intersensory facilitation of reaction time. In A. F. Sanders (Ed.), *Attention and performance III* (pp. 21-35). Amsterdam: Elsevier, North-Holland.
- COHEN, J. D., MACWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instruments, & Computers*, **25**, 257-271.
- DIEDERICH, A., & COLONIUS, H. (1987). Intersensory facilitation in the motor component? A reaction time analysis. *Psychological Research*, **49**, 23-29.
- EGETH, H. E., & DAGENBACH, D. (1991). Parallel versus serial processing in visual search: Further evidence from subadditive effects of a visual quality. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 550-559.
- ERIKSEN, C. W., GOETTL, B., ST. JAMES, J. D., & FOURNIER, L. R. (1989). Processing redundant signals: Coactivation, divided attention, or what? *Perception & Psychophysics*, **45**, 356-370.
- FOURNIER, L. R., & ERIKSEN, C. W. (1990). Coactivation in the perception of redundant targets. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 538-550.
- GIELEN, S. C. A. M., SCHMIDT, R. A., & VAN DEN HEUVEL, P. J. M. (1983). On the nature of intersensory facilitation of reaction time. *Perception & Psychophysics*, **34**, 161-168.
- GIRAY, M., & ULRICH, R. (1993). Motor coactivation revealed by re-

- response force in divided and focused attention. *Journal of Experimental Psychology: Human Perception & Performance*, **6**, 1278-1291.
- GRICE, G. R., CANHAM, L., & BOROUGHS, J. M. (1984). Combination rule for redundant information in reaction time tasks with divided attention. *Perception & Psychophysics*, **35**, 451-463.
- GRICE, G. R., CANHAM, L., & GWYNNE, J. W. (1984). Absence of a redundant-signals effect in a reaction time task with divided attention. *Perception & Psychophysics*, **36**, 565-570.
- HERSHENSON, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology*, **63**, 289-293.
- HUGHES, H. C., REUTER-LORENZ, P. A., NOZAWA, G., & FENDRICH, R. (1994). Visual-auditory interactions in sensorimotor processing: Saccades versus manual responses. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 131-153.
- MILLER, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, **14**, 247-279.
- MILLER, J. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics*, **40**, 331-343.
- MORDKOFF, J. T. (1998). Between-dimension flanker effects: A clarification with encouraging implications. *Psychonomic Bulletin & Review*, **5**, 670-675.
- MORDKOFF, J. T., MILLER, J., & ROCH, A.-C. (1996). Absence of coactivation in the motor component: Evidence from psychophysical measure of target detection. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 25-41.
- MORDKOFF, J. T., & YANTIS, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 520-538.
- NICKERSON, R. S. (1970). The effect of preceding and following auditory stimuli on response times to visual stimuli. In A. F. Sanders (Ed.), *Attention & performance III* (pp. 5-20). Amsterdam: Elsevier, North-Holland.
- NICKERSON, R. S. (1973). Intersensory facilitation of reaction time: Energy summation or preparation enhancement? *Psychological Review*, **80**, 489-509.
- NOZAWA, G., REUTER-LORENZ, P. A., & HUGHES, H. C. (1994). Parallel and serial processes in the human oculomotor system: Bimodal integration and express saccades. *Biological Cybernetics*, **72**, 19-34.
- QUINLAN, P. T., & HILL, N. I. (1999). Sequential effects in rudimentary auditory and visual tasks. *Perception & Psychophysics*, **61**, 375-384.
- RAAB, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, **24**, 574-590.
- SCHWEICKERT, R., & TOWNSEND, J. T. (1989). A trichotomy: Interactions of factors prolonging sequential and concurrent mental processes in stochastic discrete mental (PERT) networks. *Journal of Mathematical Psychology*, **33**, 328, 347.
- SPENCE, C., & DRIVER, J. (1997). On measuring selective attention to an expected sensory modality. *Perception & Psychophysics*, **59**, 389-403.
- STEIN, B., & MEREDITH, M. A. (1990). Multisensory integration: Neural and behavioral solutions for dealing with stimuli from different sensory modalities. *Annals of the New York Academy of Sciences*, **608**, 51-65.
- STEIN, B., & MEREDITH, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.
- STEIN, B., MEREDITH, M. A., HONEYCUTT, W. S., & MCDADE, L. (1989). Behavioural indices of multisensory integration: Orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience*, **1**, 12-24.
- STERNBERG, S. (1969). The discovery of processing stages: Extension of Donder's method. In W. G. Koster (Ed.), *Attention and performance II* (pp. 276-315). Amsterdam: Elsevier, North-Holland.
- TOWNSEND, J. T., & ASHBY, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.
- TOWNSEND, J. T., & NOZAWA, G. (1988, November). *Strong evidence for parallel processing with simple dot stimuli*. Paper presented at 29th Annual Meeting of the Psychonomic Society, Chicago.
- TOWNSEND, J. T., & NOZAWA, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, **39**, 321-359.
- TOWNSEND, J. T., & NOZAWA, G. (1997). Serial exhaustive models can violate the race model inequality: Implications for architecture and capacity. *Psychological Review*, **104**, 595-602.
- TOWNSEND, J. T., & SCHWEICKERT, R. (1989). Toward the trichotomy method of reaction times: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, **33**, 309-327.
- WENGER, J. W., & TOWNSEND, J. T. (2000). Basic response time tools for studying general processing capacity in attention, perception and cognition. *Journal of General Psychology*, **127**, 67-99.

NOTES

1. Although the first published account of the double-factorial paradigm is that by Townsend and Nozawa (1995), it was actually introduced in an oral presentation to the Psychonomic Society in 1988 (Townsend & Nozawa, 1988). We thank James Townsend for pointing this out to us.
2. The arguments advanced by Townsend and colleagues were based on ideas extrapolated from deterministic systems to offer mathematical proofs for the in-principle characteristics of probabilistic systems.
3. As has been pointed out in review, the tasks from which Townsend and Nozawa (1995) derived their arguments were quite different from those from which Egeth and Dagenbach (1991) derived theirs. Egeth and Dagenbach described a task in which two letters were always presented on a given trial: Participants had to decide whether a particular target letter was present. On target-present trials, a target could appear at either or both of the stimulus locations. As a consequence, target-present responses could be made on a serial self-terminating check of the locations, but target-absent responses could be made only after an exhaustive check of both locations. In contrast, Townsend and Nozawa (1995) discussed a task in which participants made yes/no responses on the basis of whether one or two dots were present in a given display. On target-absent trials, the displays were empty; they contained no dots. In this respect, a simple detection mechanism may have sufficed.
4. Again, it is important to draw a distinction between this line of reasoning based on a system incorporating deterministic principles of operation and those, discussed by Townsend and colleagues, relating to systems incorporating stochastic principles. Assumptions based on probabilistic variation in the finishing times of the various processing channels give rise to predictions that can be used to differentiate between these different kinds of systems, even though they may appear to be equivalent at a superficial level.
5. Mordkoff and Yantis (1991) have demonstrated that participants can pick up on certain stimulus-response contingencies that can influence performance under go/no-go conditions. For instance, responses on redundant target trials can be speeded when the probability of a redundant target display is high. In the present experiment, although there were strong contingencies to respond, given that eight out of nine trials demanded a response, there was no obvious way in which sensitivity to such contingencies could differentially affect performance across the four different types of bimodal trials.
6. For a thorough description of these concepts and a complete account of how to derive these various measures from data, see Wenger and Townsend (2000). We thank James Townsend for suggesting this approach to data analysis and providing reference to the cited work.

(Manuscript received August 4, 2000;
revision accepted for publication June 3, 2003.)