Running head: Distributions of Serial and Parallel Processes

The resurrection of Tweedledum and Tweedledee: Bimodality cannot distinguish serial and parallel processes

Paul Williams & Ami Eidels, University of Newcastle

James T. Townsend, Indiana University

Word count:

Main text (not including abstract, references, and footnotes) = 4,330 words
5 figures

Corresponding author:

Ami Eidels
School of Psychology
University of Newcastle
Callaghan, NSW, 2308
Australia
Emails: Ami.Eidels@newcastle.edu.au

Abstract

Simultaneously presented signals may be processed in serial or in parallel. One potentially valuable indicant of a system's characteristics may be the appearance of multi-modality in response time distributions. It is known that standard serial models can predict multimodal RT distributions, but it is unknown whether multimodality is diagnostic of serial systems, or whether alternative architectures, such as parallel, can make such predictions. We demonstrate via simulations that a multimodal reaction time distribution is not sufficient by itself to rule out parallel self-terminating processing, even with limited trial numbers. These predictions are discussed within the context of recent data indicating the existence of multi-modal distributions in visual search.

[108 words]

Key words:

Parallel, Serial, Response Times, Bimodality

Commuting to work on a busy road, the driver must process a number of signals, such as the color of the coming light and the presence (or absence) of pedestrians. Researchers often employ behavioral measures, such as response times, in an attempt to determine whether several sources of information are processed at the same time (in parallel), or one after the other (serially). Such questions were the foundation of perception and elementary cognition in the late 19th century. After lying dormant for several decades, the 1960's witnessed a renaissance of interest in human information processing including the issue of serial versus parallel search (also referred to as the "architecture issue") in memory (Sternberg, 1966) and visual displays (Egeth, 1966).

Visual search tasks have played an important role in the architecture issue. In such tasks, observers are required to indicate whether a target item is present within an array of distractors. Typically, the target item will be present on half of the trials. The number of items in the display -- set size -- is systematically varied, and researchers often plot the time it takes the observer to detect the target as a function of set size. Mean response times (RTs) typically increase with set size in a linear fashion, which suggested to researchers that search was serial in nature. However, it was soon demonstrated that parallel models could perfectly mimic these serial models and indeed, serial models could also mimic parallel models in search paradigms (e.g., Townsend, 1972; Atkinson, Holmgren & Juola, 1969; see reviews in Townsend & Ashby, 1983; Townsend, 1990; Townsend & Wenger, 2004). Over the years, new experimental paradigms and related quantitative methodologies have been developed that test parallel versus serial processing in ways that avoid the model mimicking dilemma (e.g., Ashby & Townsend, 1980; Dzhafarov & Schweickert, 1995; Schweickert, 1978; Schweickert & Townsend, 1989; Townsend, 1984; Townsend & Wenger, 2004). These methodologies commonly use signatures in the entire response time distribution to distinguish between serial and parallel processing.

An unexplored signature of RT distributions with regard to the architecture issue is that of multi-modality (a distribution that has two or more distinct maxima). In visual search, Cousineau and Shiffrin (2004) employed target and distractor stimuli that were difficult to distinguish and provided clear empirical evidence for multi-modality. On target present trials for example, they found several participants showed bimodal RT distributions for a set size of two (see Figure 2). To capture these results, they formulated a standard serial model, as well as a modified standard serial model, both of which were shown to evince multi-modality effects. The modified serial model provided the best fits. Although the parallel versus serial architecture issue was not the one they were trying to settle (their focus was on the termination issue), it is a logical step to question whether parallel models can evoke multi-modality. Because serial models are in some ways more general than are parallel models (e.g., Townsend, 1976a; cf. Townsend, 1976b; Marley & Colonius, 1992), it is unknown whether parallel models can mimic that type of behavior, or whether multi-modal distributions are indeed diagnostic of serial processing.

The question of whether *any* parallel model can predict multi-modality can be quickly cleared up. Since the first model employed by Cousineau and Shiffrin (2004) was a standard serial model, we can appeal to the conditions found by Townsend (1976a) to allow perfect parallel mimicking of a serial model. These conditions demonstrate that whenever a standard serial model predicts multi-modality, so does its perfect parallel mimicker (see Townsend, 1976a). However, parallel models that perfectly mimic standard serial models call for reallocation of resources across processing stages (e.g., Townsend & Ashby, 1983, pp. 83-91), a property that is incommensurate with parallel models that assume *independent* (i.e., non-interacting) channels. This raises our question of interest: can a current independent parallel model from the literature *plausibly* produce multi-modality?

We begin by elaborating on the seminal example of serial-parallel mimicry in visual search tasks, where observing an increase in mean RT as set-size increases was typically taken to suggest a serial search strategy. We then move on to consider the limitations of this interpretation, recapping the work of Townsend and others. Next, we outline the computer simulations used to show another case of parallel-serial mimicry, where a plausible independent channel self-terminating parallel system can produce bimodal RT distributions. We describe the basic setup of the simulations and the assumptions of the model, which utilize a common model of decision-making, the linear ballistic accumulator (LBA; Brown & Heathcote, 2008). Lastly, we show that a commonly available statistical test can detect this bimodality with as few as 1,000 simulated trials – a reasonable experimental sample[1]. We conclude that empirically observing a bimodal RT distribution is insufficient to reject accounts based on parallel processing, even certain instances of independent parallel processing.

Visual search and parallel-serial mimicry

In a typical visual search task (e.g., Treisman & Gelade, 1980) observers are requested to detect a target-item among distractors as quickly as possible. The target may differ from the distractor item(s) by one feature (say, colour -- a blue target letter among red distractors), or require a combination of features to distinguish it from the distractor items (say, colour and form -- a blue X among red Xs, red Os, and blue Os). Another important manipulation is set size; trials may contain a different number of items in the display, sometimes including a target item (positive trials) and sometimes not (negative trials). The inclusion of negative trials is important so that participants cannot repeatedly press the "yes" button. Example displays for a single-feature search and two-feature search are presented in Figure 1a and 1b, respectively.

[1] 1,000 is the same sample size reported in Cousineau and Shiffrin (2004).

Hypothetical results from such an experiment (e.g., Treisman & Gelade, 1980) are presented in Figure 1c. For the single-feature condition, in this instance finding a blue letter among red letters, typical search latencies are roughly the same regardless of display size. That is, the function relating mean RT as a function of the number of items in the display is flat with a slope of zero. This finding had led to the notion that single-feature search is carried out in parallel (Treisman 1988, 1992). Search for a unique combination of features (conjunction), such as blue X among distractors that can share with the target either form or colour, resulted in a monotonically increasing function. This result had been interpreted as the outcome of a slower, serial search that requires attention.

However, this parallel-serial interpretation is not altogether complete. As Townsend and others have shown, there are conditions under which parallel and serial models are mathematically equivalent (e.g., Townsend & Ashby, 1983; Townsend, 1990; Schweickert, 1978; Dzhfarov & Schweickert, 1995). More bluntly put, a parallel model can predict a monotonically increasing mean RT function, and even the straight line prediction of standard serial models[2], as set size increases (Algom, Eidels, Hawkins, Jefferson, & Townsend, 2013). Empirically observing a monotonically-increasing RT function in a search task then, even a linear function, is insufficient to determine whether processing multiple items is carried out in serial or in parallel. To overcome this mimicking problem, researchers had to develop new methodologies (e.g., Townsend & Nozawa's Systems Factorial Technology, 1995) or look for other 'empirical signatures' that are unique to certain models. One such empirical signature that might be considered a typical prediction of serial processes is the bimodal (or multimodal) RT distribution.

[2] Parallel processes with *unlimited capacity* (average time to process one item does not vary with the overall number of item presented and processed) predict a negatively accelerating mean RT vs. set-size function. However under certain assumptions, *limited-capacity* parallel models can mimic the linear function that is predicted by serial models (Townsend, 1976a).

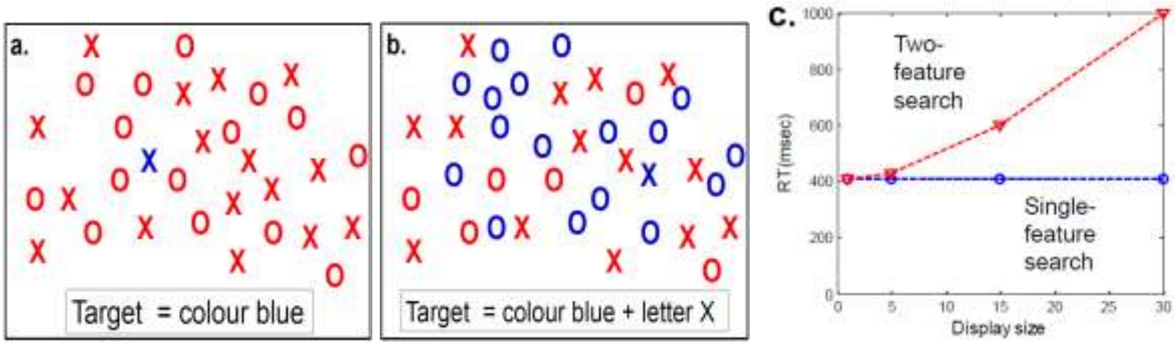

Figure 1. Example displays in a visual search task. Target detection in Panel a requires a single feature search, whereas Panel b requires two features. Panel c shows typical results.

The case for bimodal RT distribution in serial and parallel processes

A serial mode of processing makes the following testable prediction: the number of items that must be scanned before finding the target and responding should vary across trials, depending on where observers start their search and the position of the target. The consequence is a mixture of fast responses from when the target was detected early in the search, and slow responses from when it was detected later on. Combining together distributions of faster or slower responses, each having a different mode, results in a single mixture distribution that may be bimodal (or multimodal, if more than two positions are involved; for simplicity we focus on two processes from now on). The question then is about the interpretation of bimodality when it is observed: does a bimodal RT distribution necessarily mean that the underlying processes are serial?

As discussed earlier, Cousineau and Shiffrin (2004) provided one of the most compelling examples of multimodality in empirical data. They reported data from three participants engaged in a visual search task, where the target and distractor stimuli were constructed to encourage serial search (the items were wheels with spokes and the target was defined by a conjunction of specific spokes). Re-examining their data, Donkin and Shiffrin

(2011) noted that “positive responses [in Cousineau & Shiffrin 2004 data] exhibited *multimodal* response time distributions, with the modes roughly corresponding to the serial position in which the target happened to be compared.” (p. 2, emphasis added). Although Cousineau and Shiffrin were primarily interested in the termination rule (whether participants scan the entire visual field, or stop at some earlier point such as identifying a target), their data provide a clear demonstration of multimodality in RT distributions from a visual-search task. Example data from three participants, with display sizes (DS) of 1, 2, and 4 items are presented in the left panel of Figure 2.

Freeman and Dale (2013) extended the scrutiny of unimodal versus bimodal distributions to another dependent variable, arm-reaching trajectories. They stipulated that “…two processes that work on a different temporal scale… predict that the distribution of behavioral measures based on these responses will exhibit *bimodality*.” (p. 84). The right panel of Figure 2 illustrates trials from a hypothetical experiment where responses were given by moving a computer mouse to the right or left side. A mixture of trials from two different kinds (e.g., straight vs. convex trajectories) results in bimodal distribution of the dependent variable (distance along the x-axis, in this example). It is crucial to note at this point that the two processes need not occur sequentially, one after the other, to exhibit bimodality. Rather, they can start at the same time, one being slower than the other. This situation can be readily accomplished by an independent parallel system that processes one item quickly and efficiently, and the other more slowly. This idea forms the rationale for our simulations.

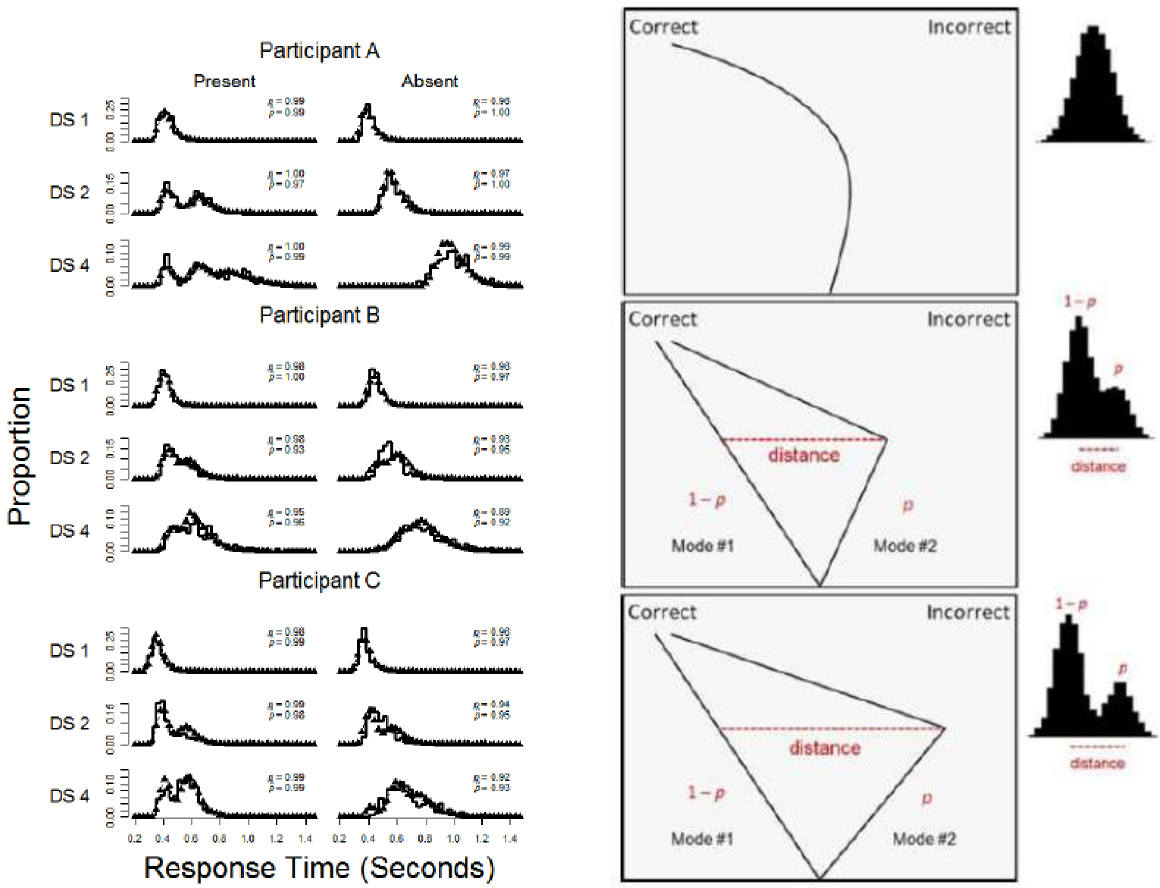

Figure 2. Example data from Cousineau & Shiffrin, 2004 (left) and from Freeman & Dale, 2013 (right).

Constructing a parallel-model simulation that produces a bimodal RT distribution

The observable RT density function of a two-state model may be bimodal as it is a probability mixture of two unimodal distributions. A serial model with two stages of processing is one instance of such a model. In visual search the stages may reflect the order of processing, with position *a* processed first and position *b* processed second. However, a parallel model with fast and slow processes for positions *a* and *b*, respectively, also qualifies as “mixing two distinct types of response” (Townsend & Ashby, 1983, p. 263). Next, we describe an example of such model in the context of visual search, and demonstrate that it can reasonably generate a bimodal RT distribution.

Our simulation is set up to mimic a visual search experiment, with a fixed set size of two. Thus, on each simulated trial there are two visual positions (e.g., two horizontally aligned items, one to the left and the other to the right of a centrally positioned fixation-point). While we are only interested in the results of trials in which the target is present, our simulation works like a typical visual search experiment. Thus, on half of the trials the target is present, whereas for the other half the target is absent. When present, the target is positioned randomly (with an equal overall number for each position).

The model: parallel processing with attention gradient

We tested a simple parallel LBA model with ‘present’ and ‘absent’ accumulators for both the left and right position. Each accumulator collects evidence towards some prescribed threshold, at some rate (as described below). If either target-present accumulator reaches threshold before its corresponding target-absent accumulator, a target-present response is triggered immediately. This represents self-terminating search. For an absent response to be triggered, both target-absent accumulators must reach threshold before their corresponding target-present accumulators. This represents exhaustive search. As soon as a response is triggered the trial is terminated.

Visual attention may affect processing efficiency by way of improving the quality of evidence collection (increasing the rate of accumulation) in the attended region, by reducing criterion in that area, or both. Regardless of the exact mechanism, it is said to have salutary effects on performance, as suggested, for example, by Downing and Pinker (1985) and more recently by Müller, Mollenhauer, Rösler, and Kleinschmid (2005). Downing and Pinker proposed a Gaussian attention gradient that enhances processing of visual stimuli within a circumscribed region of space. Alternately, Muller et al. proposed a Mexican Hat function of attention modulation (see also Hopf, Boehler, Luck, Tsotsos, Heinze & Schoenfeld, 2006;

Carrasco, 2011, for a review). The exact form of the modulation function is not important for our efforts, and may actually vary depending on task difficulty. The critical point is the advantage in processing for items that are closer to the centre of attention.

We assumed that attention is always directed to the left position, so items on that side fall within focal attention whereas items on the opposite side get less attention. This assumption can be relaxed without loss of generality, as long as the location of attention and the location of the target item remain unrelated. For simplicity, we kept attention at a fixed location. The attentional advantage can be modelled by a number of model parameters, depending on the accumulator model chosen for the task. Our model of choice was the LBA, which we describe next.

The linear ballistic accumulator

The LBA (Brown & Heathcote, 2008) is a model of rapid choice that has 5 parameters: basetime (non-decision time), maximum possible start-point, mean drift rate, drift-rate variance, and response threshold. A response in the LBA model begins with a random amount of evidence, a, sampled from a uniform distribution between 0 and A. Evidence accumulates at a linear and fixed (i.e., ballistic) rate that is sampled from a normal distribution with mean, v, and standard deviation, s, until a threshold amount, b, is collected. The observed response time is the sum of the time taken for evidence to reach threshold in the accumulator associated with the chosen response plus the time taken for non-decision aspects of response time, such as the time taken to encode stimuli and the time taken to execute the motor response, t_0. Formally, response time on each trial can be written as $T = t_0 + (b-a)/v$.

Because evidence is accumulated linearly and at a fixed rate, there is no within-trial variability (unlike other successful models of choice and RT). This simplification allowed

Brown and Heathcote (2008) to derive closed-form solutions for the probability density function, f(*t*), and the cumulative distribution function, F(*t*), and makes the LBA easy to simulate.

The set-up for the simulation is illustrated at the bottom of Figure 3. There are four parallel-independent linear and ballistic accumulators - present and absent accumulators for each position. As mentioned before, the response on each trial is determined by the outcomes in each position. If either target-present accumulator reaches threshold before its corresponding target-absent accumulator, a target-present response is triggered immediately. In contrast, a target-absent response is triggered only if both target-absent accumulators reach threshold before their corresponding target-present accumulators. Figure 3 illustrates a two-item display, with distractor O on the left and target X on the right. Since we assume that the left position is within focal attention (illustrated by the yellow circle), the accumulators corresponding to the left item should collect evidence at a higher rate and/or require less evidence (lower threshold).

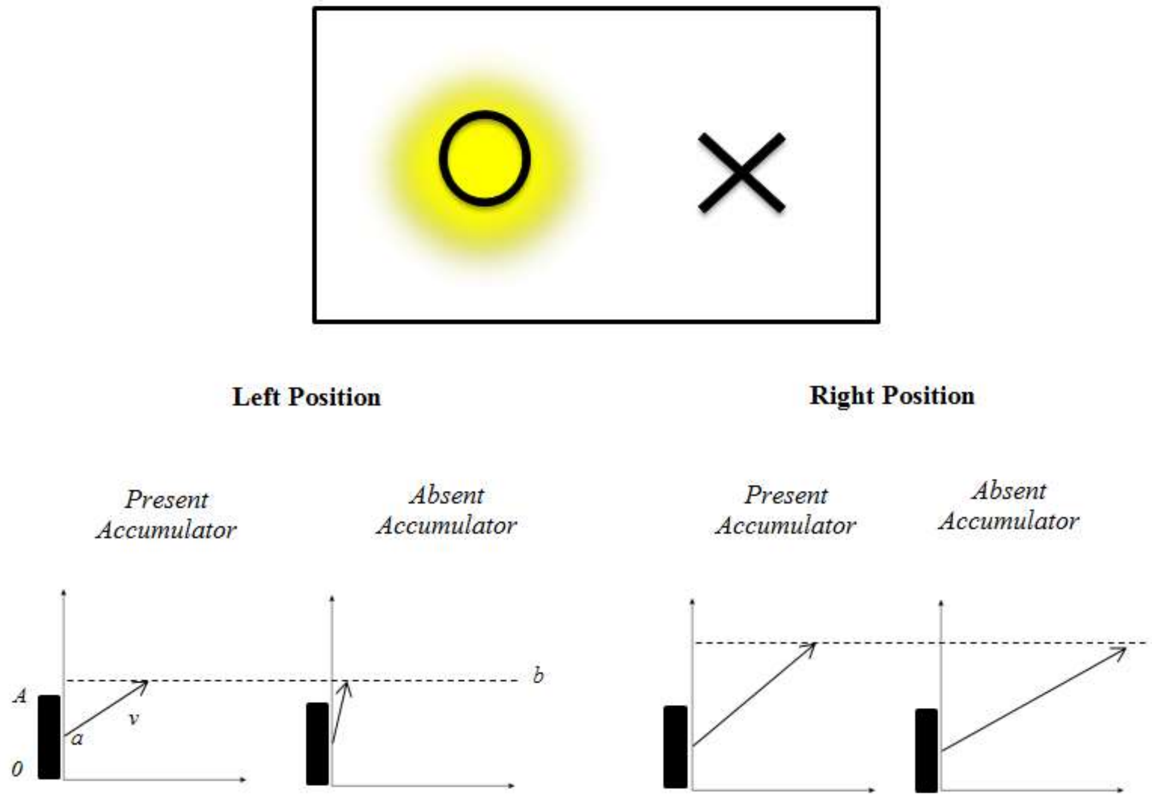

Figure 3. Illustration of the four parallel-independent accumulators set-up for the simulation. The top box illustrates a possible display in a hypothetical visual-search task. The bottom illustrates the corresponding model. The more-attended region on the left of the display is marked by the yellow circle. Since the target-item, X, is on the less-attended side of the display, its corresponding correct accumulator (target-present) has a lower drift rate than the correct accumulator on the left (target-absent). The threshold is also displayed as lower for the more-attended accumulators. Note *a* has been standardized across all accumulators here for simplicity, but in simulation is drawn randomly from a uniform distribution between 0-*A* on each trial for each accumulator. Also note *s* and t_o are essential parameters of the LBA, but are not represented in this figure (see text for detail).

The simulation

Simulation details

We used two accumulators for the left, more-attended position, and two for the right, less-attended position. For the correct accumulator in the more-attended position (in Figure 3, the target-absent accumulator on the left), we fixed the values of all five parameters as follows: s=1, t_0=0.1, A=0.5, b=1, and v=6. These values are approximately based on estimates we have obtained from other studies using the LBA to fit empirical data (e.g., Eidels, Donkin, Brown, & Heathcote, 2010; Eidels, 2012). For the less-attended correct accumulator (in Figure 3, the target present accumulator on the right), we fixed the values of s, t_0, and A to be the same as those of the attended accumulator. Critically, we systematically varied the values of the rate and threshold parameters of the less-attended correct accumulator. We allowed the drift rate, v, to vary from 6 to 0 (in steps of -.05). We similarly varied the values of the threshold parameter, b, from 1 to 2 (in steps of .01). For both levels of attention, the only parameter to differ between correct accumulators and their incorrect counterparts was the drift rate, which was reduced by a factor of 1/2.

Overall, our systematic changes resulted in 12,220 combinations of rate and threshold (in the less-attended accumulator). We ran two simulations. In the first, we generated 100,000 target present trials (of 200,000 total trials) for each rate and threshold combination. This simulation demonstrates that bimodality is achievable with plausible LBA parameter estimates. In the second, we generated only 1,000 target present trials for each rate and threshold combination - the same amount reported by Cousineau & Shiffrin (2004). We assessed the correct responses for multi-modality using Hartigan's dip test (Hartigan & Hartigan, 1985), a commonly available statistical test[3]. This simulation demonstrates that bimodality can be achieved with a plausible number of experimental trials.

Results

Figure 4 shows results from our first simulation for a selected subset of parameter values. The RT distributions displayed are for correct responses to target-present trials. Each column in the figure corresponds to a different combination of parameter values. Drift rate

[3] Freeman & Dale (2013) compared three tests for bimodality: bimodality coefficient, Hartigan's dip statistic, and the Akaike information criterion difference (AIC_{diff}). Although the AIC_{diff} test had the highest rate of hits (95%), it also had the highest false-alarm rate (81%), where false alarm is detection of bimodality when the true model is unimodal. The dip test had the highest d' (3.43; two times that of the bimodality coefficient, four times the AIC). It is also more robust to skew, so is the most suitable for RT distributions.

for the more-attended accumulator was fixed at 6, but the rate for the less-attended accumulator was 5 (left column), 4.5 (second from left), 4 (third from left), and 3.5 (right-most column). The threshold for the attended and unattended correct accumulators in this example were 1 and 1.4, respectively. The rows correspond to the source of the distribution. The top row presents RT distributions for the attended accumulator. The middle row presents RT distributions for the less attended accumulator. The bottom row presents the distribution of the mixture of the two previous distributions. Note that the first two rows of distributions are not available for the scientist to view in empirical data, since trials from the two positions are intermixed. It is only the empirical distribution, at the bottom of the figure that is available for the researcher to inspect. In particular, the bottom row shows how bimodality builds up in the observed distribution as we decrease the rate of evidence accumulation in the less-attended channel, from left to right. Error rates also increase from left to right (5.18%, 6.04%, 7.17%, and 8.35%, respectively).

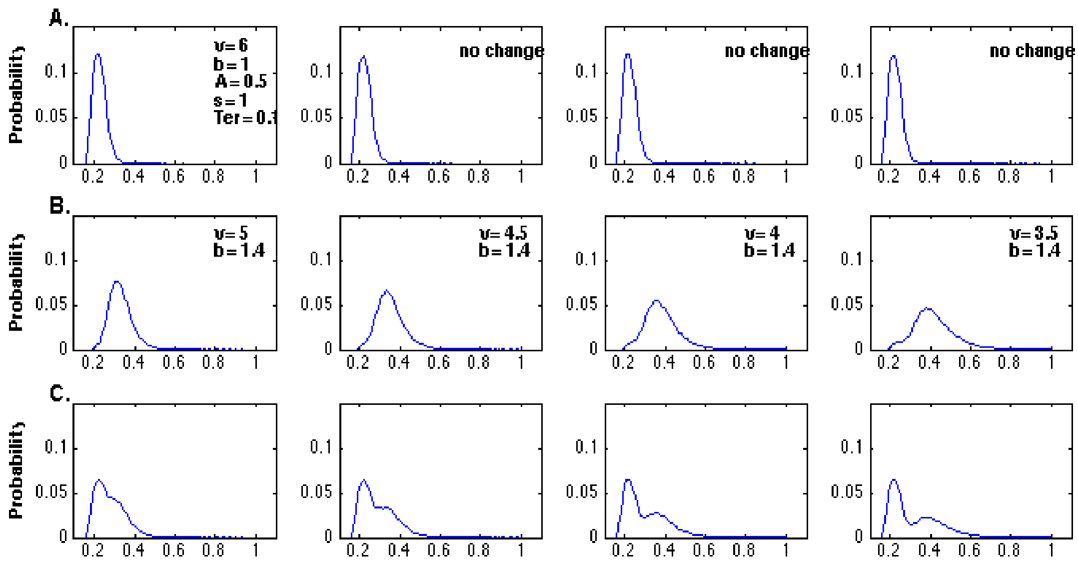

Figure 4. Simulation results for correct target-present decisions for a selected subset of parameter values. RT distributions for targets in the attended position (top row) and less attended position (middle row). The bottom row shows the (empirical) mixture distributions. See text for details.

The results depicted in Figure 4 demonstrate that an independent (self-terminating) parallel system can produce bimodal distributions given plausible parameter estimates. Importantly, Figure 4 also shows that the RT's in our simulations are realistic. In our simulation, the target item appeared in the less-attended region on half of the trials. For this half of the trials then, the unattended accumulator determines RTs. As lowering drift rate and/or increasing threshold leads to slower RTs, it is possible that bimodality could be detected only by mixing very slow trials from this unattended accumulator - slower than what is typically observed in empirical data - with faster RTs determined by the attended accumulator. Figure 4 suggests that this is not the case.

Figure 5 shows the results of Hartigan's dip tests on each combination within the entire parameter space for both simulations. To recount, the simulations differ only in the number of trials (100'000 vs. 1000 target present-trials). The horizontal x and y axes of the figure thus represent the systematic changes in the drift rate (*v*) and threshold (*b*) parameters of the less-attended accumulator. The vertical axis shows the *p* values of the Hartigan's dip test, with small values ($p<.05$) rejecting the null hypothesis that the observed distribution is unimodal. The larger simulation (left panel) shows a clear diagonal divide highlighting the space in which bimodality can be detected, while the smaller simulation (right panel) shows a noisier pattern. This is commensurate with the drop in power expected given the reduction in trial numbers by a factor of 100. Crucially, it is evident from both panels that bimodal RT distributions are not a unique prediction of serial processes or even of serial-mimicking parallel systems; an independent self-terminating parallel system with the justifiable addition of attention gradient can lead to bimodal RT distributions – even for a substantial part of the parameter space using only 1000 trials.

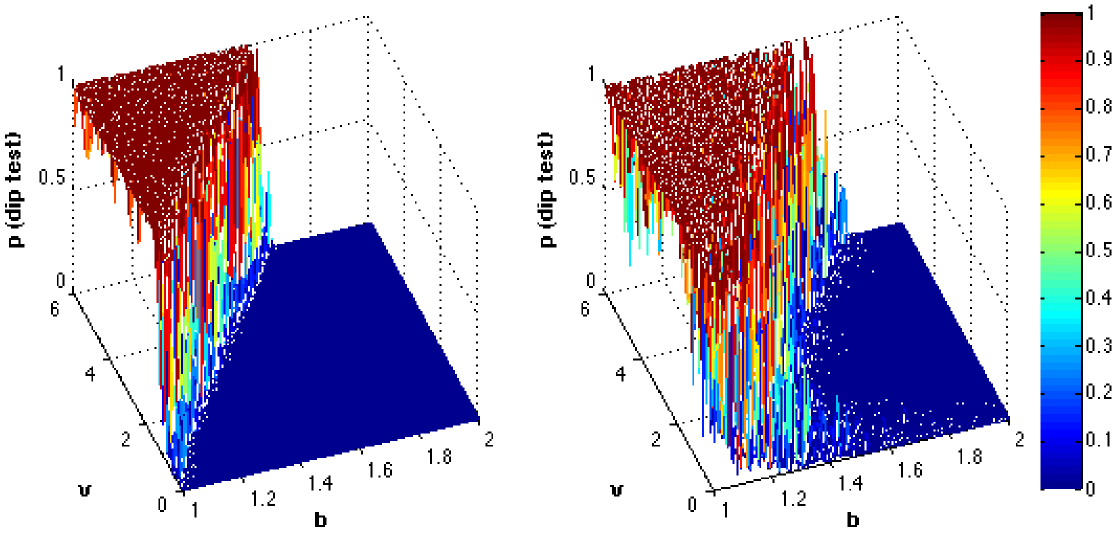

Figure 5. *p* values of the Hartigan's dip test for bimodality as a function of the rate (*v*) and threshold (*b*) in the less-attended accumulator. The left panel presents the simulated results for 100,000 target-present trials, while the right panel displays the (noisier) results for 1,000 target-present trials. Small *p* values suggest rejecting the null hypothesis of a unimodal RT distribution.

General Discussion

Our simulation results, summarized in Figure 5, suggest that an independent parallel model with the addition of an attention gradient can generate a bimodal response-time distribution that is detected by the Hartigan's dip test with as few as 1,000 trials. Bimodality can be achieved by increasing the decision threshold in less-attended region(s), by lowering the evidence accumulation rate in the less-attended region, or by various combinations of the two. Bimodality is therefore insufficient to rule out self-terminating parallel models, even those with independent channels.

These results are commensurate with earlier studies on the effects of visual attention using signal detection theory. Farah (1989, Experiment 4) used signal-detection measures that separate perceptual and decisional factors, and showed that allocating visual attention to a particular region in space reduced criterion (beta) in that region. More recent studies, using a different method, debate whether attention affects the decision threshold or the quality of the percept (i.e., evidence accumulation rate). For example, Carrasco and colleagues (e.g., Carrasco & McElree, 2001; Anton-Erxleben, Abrams, & Carrasco, 2010) argued that attention enhances rate of evidence accumulation and apparent contrast but does not affect bias, whereas Schneider (e.g., Schneider, 2011; Schneider & Komlos, 2008) has stressed that attention alters decision threshold but not appearance. Our study takes no side in this debate and makes no claim about how attention enhances performance. Rather, it simply aims to show that bimodality could result from effects of such an attention gradient, whether improvement is in rate, criterion, or both.

Close inspection of Figure 5 may suggest that the attentional effect of threshold change on bimodality is more robust than that of rate; increasing threshold by 60% from 1 to 1.6 (while rate was held fixed) resulted in detection of bimodality by the dip test (at p<.05), whereas a similar result based on rate changes required that rate decreased by half an order of magnitude, from 6 to about 1.2 (while the threshold was fixed). This outcome, however, is likely to depend on the parameters of the simulated model. In particular, the maximum value of the start point distribution, A, was initially set to 0.5. Recall the amount of evidence to accumulate is given by subtracting the start point, a (sampled from a uniform distribution ~U(0, A)), from the threshold, b. With an initial threshold of 1, relatively small changes to b will result in proportionally dramatic changes of the expression b-a. However, with a smaller value of A, the same changes in the decision threshold should have less impact. We tested this by repeating the same simulations for various values of A and found when A was set to 0.2, the effects of rate and threshold were roughly equal (i.e., a similar proportional change in either rate or threshold was required to obtain confidence in the bimodality of the RT distribution).

As always, there are limitations to the current arguments. For instance, we do not profess that parallel processing is likely under every experimental design. For instance, by making the targets difficult to identify, Cousineau and Shiffrin (2004) increased the likelihood of serial processing. One can even envisage an experimental design that employs the use of eye tracking so that the target appears only upon fixation. This would make parallel processing impossible – even after unlimited practice. Lastly, it is important to note that we do not profess that the underlying architecture of our independent self-terminating parallel model naturally gives rise to bimodal distributions. Rather, it is the justifiable addition of an attention gradient that makes bimodality possible.

Conclusions

Multi-modality might seem like a trademark of serial systems. While no analytic results exist for general or parameterized serial models as to the conditions that suffice to elicit bimodality, the models constructed elsewhere and successfully applied to multi-modal data have been based on serial architectures - which forms a type of existence proof. The situation for parallel systems has been even murkier. Not only was there no analytic work carried out on such systems to date, it has not been made manifest that they are even capable of producing multi-modal distributions (though Cousineau & Shiffrin (2004) allow for that possibility).

We therefore began by recapping the general serial-parallel equivalence theorem that demonstrates parallel models can, a fortiori, mimic standard serial models. However, these parallel models that mimic standard serial models are a special case. Unlike so-called standard parallel models they allow some form of channel interaction, by reallocating resources or capacity across processing completion stages (e.g., see Townsend & Ashby, 1983, pp. 48, 88, or 138). Standard parallel models, on the other hand, assume independent channels. Thus, we were interested as to whether self-terminating independent parallel models could be capable of predicting multi-modal processing time distributions. Based on our results, we conclude that observing a bimodal RT distribution in an experiment (such as, but not restricted to, visual search) is not sufficient to reject accounts that are based on either serial-mimicking parallel system, or even standard, self-terminating, independent-channels parallel processing.

References

Algom, D., Eidels, A., Hawkins, R. D., Jefferson, B., & Townsend, J. T. (2013). Features of response times: Identification of cognitive mechanisms through mathematical modeling. To appear in J. R. Busemeyer, J. T. Townsend, Z. J. Wang, & A. Eidels (Eds.), *Oxford Handbook of Computational and Mathematical Psychology*, Oxford University Press.

Anton-Erxleben, K., Abrams, J., & Carrasco, M. (2010). Evaluating comparative and equality judgments in contrast perception: Attention alters appearance. *Journal of Vision, 10*, 1-22.

Ashby, F. G., & Townsend, J. T. (1980). Decomposing the reaction time distribution: Pure insertion and selective influence revisited. *Journal of Mathematical Psychology, 21,* 93-123.

Atkinson, R. C., Holmgren, J., & Juola, J. F. (1969). Processing time as influenced by the number of elements in a visual display. *Perception & Psychophysics*, *6*, 321-326.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*, 153-178.

Busemeyer, J., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making. *Psychological Review, 100,* 432-459.

Carrasco, M. (2011). Visual Attention: The past 25 years. *Vision Research*, *51,* 1484-1525.

Carrasco, M. & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proceedings of the National Academy of Sciences*, *98*, 5363-5367.

Cousineau, D., & Shiffrin, R. M. (2004). Termination of a visual search with large display size effects. *Spatial Vision*, *17,* 327-352.

Donkin, C. & Shiffrin, R. (2011). Visual search as a combination of automatic and attentive processes. In L. Carlson, C. Hoelscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Marin (Eds.). *Attention and Performance (Vol. XI).* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Dzhafarov, E. N., & Schweickert, R. (1995). Decompositions of response times: An almost general theory. *Journal of Mathematical Psychology*, *39*, 285-314.

Egeth, H. E. (1966). Parallel versus serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics*, *1*, 245-252.

Eidels, A. (2012). Independent race of colour and word can predict the Stroop effect. *Australian Journal of Psychology*, *64*, 189-198.

Eidels, A., Donkin, C., Brown, S. D., & Heathcote, A. (2010). Converging measures of workload capacity. *Psychonomic Bulletin & Review*, *17*, 763-771.

Farah, M. J. (1989). Mechanisms of imagery-perception interaction. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 203-211.

Freeman, J. B., & Dale, R. (2013). Assessing bimodality to detect the presence of a dual cognitive process. *Behavior Research Methods, 45*, 83-97.

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.

Hartigan, J. A., & Hartigan, P. M. (1985). The dip test of unimodality. *The Annals of Statistics*, 70-84.

Hopf, J. M., Boehler, C. N., Luck, S. J., Tsotsos, J. K., Heinze, H. J., & Schoenfeld, M. A. (2006). Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 1053–1058.

Kendall, M. G., Stuart, A (1969). The advanced theory of statistics (Vol. 1). London: Griffin.

Marley, A. A. J., & Colonius, H. (1992). The “horse race” random utility model for choice probabilities and reaction times, and its compering risks interpretation. *Journal of Mathematical Psychology, 36,* 1-20.

Müller, N. G., Mollenhauer, M., Rosler, A., & Kleinschmidt, A. (2005). The attentional field has a Mexican hat distribution. *Vision Research, 45,* 1129–1137.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59-108.

Schneider, K. A. (2011). Attention alters decision criteria but not appearance: A reanalysis of Anton-Erxleben, Abrams, and Carrasco (2010). *Journal of Vision*, *11,* 1-8.

Schneider, K. A., & Komlos, M. (2008). Attention biases decisions but does not alter appearance. *Journal of Vision*, *8*, 1-10.

Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology*, *18*, 105-139.

Schweickert, R., & Townsend, J. T. (1989). A trichotomy method: Interactions of factors prolonging sequential and concurrent mental processes in stochastic PERT networks. *Journal of Mathematical Psychology, 33*, 328-347.

Smith, P. L., & Van Zandt, T. (2000). Time-dependent Poisson counter models of response latency in simple judgment. *British Journal of Mathematical and Statistical Psychology, 53,* 293-315.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*(3736), 652-654.

Townsend, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology, 25,* 168-199.

Townsend, J. T. (1976a). A stochastic theory of matching processes. *Journal of Mathematical Psychology, 14,* 1-52.

Townsend, J. T. (1976b). Serial and within-stage independent parallel model equivalence on the minimum completion time. *Journal of Mathematical Psychology, 14*, 219-238.

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology, 28,* 363-400.

Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science*, *1*, 46-54.

Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.

Townsend, J. T., & Eidels, A. (2011). Workload capacity spaces: A unified methodology for response time measures of efficiency as workload is varied. *Psychonomic Bulletin & Review*, *18*, 659-681.

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. & Wenger, M.J. (2004). The serial-parallel dilemma: A case study in a linkage of theory and method. *Psychonomic Bulletin & Review, 11*, 391-418.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology*, *40*, 201-237.

Treisman, A. (1992). Spreading suppression or feature integration? Author's reply. *Journal of Experimental Psychology. Human Perception and Performance*, *18*, 589-593.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97-136.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550-592.

Author note

We thank Andrei Teodorescu and two anonymous reviewers for helpful comments on an earlier version of this manuscript. The study was supported by an Australian Research Council Discovery grant, ARC-DP 120102907, to A. E. Please address correspondence to Ami.Eidels@newcastle.edu.au.