The influence of cueing on attentional focus in perceptual decision making

Cheng-Ta Yang · Daniel R. Little · Ching-Chun Hsu

Abstract Selective attention has been known to play an important role in decision making. In the present study, we combined a cueing paradigm with a redundant-target detection task to examine how attention affects the decision process when detecting the redundant targets. Cue validity was manipulated in two experiments. The results showed that when the cue was 50% valid in one experiment, the participants adopted a parallel self-terminating processing strategy, indicative of a diffuse attentional focus on both target locations. When the cue was 100% valid in the second experiment, all of the participants switched to a serial self-terminating processing strategy, which in our study indicated focused attention to a single target location. This study demonstrates the flexibility of the decision mechanism and highlights the importance of top-down control in selecting a decision strategy.

Keywords Attention · Decision making · System factorial technology

In this article, we examine how cue validity affects attentional focus to a target location. Though a number of parametric models of attention have been developed, we characterize attentional focus in a nonparametric fashion along two dimensions, the processing architecture and the processing capacity of an information processing system. Both of these measures provide evidence for whether attention is spread across two targets or localized to a single target when accumulating information for a detection decision. We utilize a redundant-target detection task, in which a target is presented in either one of two well-localized locations, both locations, or neither location, and an observer must respond yes if a target is detected in either location.

The task factorially combines the brightness of a target with the presence of a target at both or only one location; these two manipulations allow for the application of several systems factorial technology (SFT; Townsend & Nozawa, 1995; Townsend & Wenger, 2004) measures that can be used to diagnose the processing architecture (e.g., serial vs. parallel vs. coactive processing), the decisional stopping rule (self-terminating vs. exhaustive rule), and processing capacity (limited capacity vs. unlimited capacity vs. supercapacity). Our goal in this article is to examine how an information-processing strategy is affected by the presence of an informative or uninformative cue.

The redundant-target detection task has been widely used to investigate how information from multiple sources is processed to trigger a single response (Miller, 1982, 1986, 1991; Miniussi, Girelli, & Marzi, 1998; Mordkoff & Yantis, 1991, 1993; Townsend & Nozawa, 1995). The processing requirements of this task are well-known (see Townsend & Nozawa, 1995, for a review). In a typical version of this task, participants are required to supervise two distinct locations. When any dot stimulus at either location is detected, participants have to make a speedy response; otherwise, if no dot stimulus is detected, then an alternative response is to be emitted. Previous studies have shown that reaction times (RTs) in the double-dot condition (both locations contain a target) are faster than those in the single-dot condition (only one location contains a target). This phenomenon is called the redundant-target effect (RTE) or the redundancy gain (RG).

Two major accounts have been proposed to explain the RTE: First, according to the horse-race model or the independent-race model (Raab, 1962), the RTE is observed because of the statistical facilitation from the two different
channels. That is, the distribution of minimum completion times for processing two targets (double-dot condition) will be less than the distribution of processing times for any one target (single-dot condition); this is expected if processing both locations is accomplished in a parallel, self-terminating fashion, in which each location is processed independently but in a simultaneous fashion. Alternatively, the coactivation model suggests that the RTE is observed because information from multiple channels is integrated or pooled together into a single channel before a decision is made (Miller, 1982). Using SFT (Townsend & Eidels, 2011; Townsend & Nozawa, 1995), previous studies have shown that when the targets are presented without cues to target presence (or absence), the information processing characteristics of this detection task are well-defined and identifiable. Namely, processing is clearly parallel, self-terminating, and of limited capacity.

Although the two accounts differ regarding the integration process in decision making, both are mute to the question of whether information processing is sensitive to other experimental manipulations. These accounts can be classified broadly as indicating a type of process invariance. In contrast to this process invariance hypothesis, the relative saliency hypothesis (Yang, 2011; Yang, Chang, & Wu, 2013; Yang, Hsu, Huang, & Yeh, 2011) suggests that the detection decision process is flexible and can vary according to the relative salience or relative importance of different sources of information. When information from multiple channels is equally salient and important for detection making, parallel processing is likely to be adopted. Conversely, when one source of information is more salient or more important than the other, serial processing is adopted.

In a change detection task, if the salience of one dimension is increased by increasing the probability of a change on that dimension, processing shifts from parallel to serial (Yang et al., 2013). That is, observers switch attention solely to the dimension that is more likely to change on a given trial. Here we investigate a related manipulation by systematically varying the validity of an exogenous cue to target location in a redundant-target detection task.

On each trial, a cue was presented at one of two peripheral boxes. After the cue display, a test display was presented and participants had to detect the target dot. The validity of a cue to the target location in the single-dot conditions was manipulated in order to manipulate the utilization of the cue and map out its effect on the processing characteristics of a system. The cue was manipulated such that it was either uninformative (50 % valid) or informative (100 % valid) in two experiments. If attention is oriented to one of the two locations with higher priority, we expect that that location will be processed first, and in the case when a target is located at the cued location, exclusively (i.e., the other location will not be processed). This would be consistent with an attentional focusing to the cued location (LaBerge, 1983). In the present task, this change in processing would manifest as a change from parallel processing of both locations to serial self-terminating processing with the cued location processed before the uncued location.

A change from parallel to serial processing might only be expected to occur when the cue was of sufficient validity; by contrast, when the validity of the cue was reduced, if attention is controlled in a top-down manner, participants should then ignore the cue and maintain a parallel processing strategy. On the other hand, an alternative possibility is that attention to any onset cue is mandatory (Schreij, Owens, & Theeuwes, 2008; Theeuwes, 2004b), which would predict serial processing with any cue, regardless of validity. Finally, it may be the case that the redundancy gain offers an advantage in processing time for parallel processing of both locations, over and above any benefit gained by focusing attention serially on only one location. If this is true, then the cue may be ignored altogether, and processing may remain parallel in all cases. Here we used a methodology (SFT) that could effectively differentiate these hypotheses.

Introduction to SFT

The experiment was designed on the basis of the double factorial paradigm, which in SFT provides a basis for measuring information processing (see Townsend & Nozawa, 1995, for a review). SFT allows for diagnoses of the characteristics of a dynamic processing system, such as the processing architecture, the decisional stopping rule, and the processing capacity.

The processing architecture denotes the processing order of multiple sources of information. In the models that we consider in this article, serial processing means that each source of information is processed in sequence, without any overlap of the individual processing times. Parallel processing denotes that multiple sources are processed simultaneously. Coactive processing is similar to parallel processing, but the separate activations from multiple channels are weighted and combined into a total activation before a decision is made. Each of these processing models is illustrated in Fig. 1.

Also shown in Fig. 1 is a distinction in the amounts of information required for decision making (i.e., the decisional stopping rule). A self-terminating stopping rule means that a decision is based on the faster completed signal. When one of the signals reaches the decision criterion, a decision is made, and the processing of the other signals will be terminated. An exhaustive stopping rule means that a decision is based on all of the signals; that is, a decision is not made until all of the signals are completely processed. For instance, in the present tasks, if a target was detected in the first processed location of a serial search, then processing might self-terminate with a “target detected” decision. On the other hand, serial
processing might continue to process the remaining location in an exhaustive fashion.

The processing capacity denotes the variation of the processing efficiency as a function of the workload (the number of signals to be processed). The process is limited capacity when the average of individual processing times increases as a result of workload increases; it is unlimited capacity when the average of individual processing times is not affected by the workload; and it is supercapacity when the average of individual processing times decreases as the workload increases.

To diagnose each of the information-processing characteristics, we utilized a factorial combination of target brightness in each location in the double-dot trials of the detection task (see Fig. 2). Presumably, a high-brightness dot is detected...
faster and more accurately than a low-brightness dot (e.g., Luce, 1986). As a result, we used four types of test trials in the double-dot conditions: both dots could have high brightness (HH\(^1\)), the left dot could have high brightness and the right dot low brightness (HL), the left dot could have low brightness and the right dot high brightness (LH), or the two dots could have low brightness (LL). Also, four types of test trials were used in the single-dot conditions: the left dot could have high brightness (HX), the left dot could have low brightness (LX), the right dot could have high brightness (XH), or the right dot could have low brightness (XL). Figure 2 presents all nine possible combinations of the test trials.

Using the RTs from the HH, HL, LH, and LL conditions, we can compute the mean interaction contrast (MIC) and survivor interaction contrast (SIC), to infer the processing architecture and the decisional stopping rule. The MIC can be expressed as

\[
MIC = RT_{(H,H)} - RT_{(H,L)} - RT_{(L,H)} \cdot RT_{(L,L)},
\]

where \(RT\) indicates the mean RT in one of the double-dot conditions. The two subscripts refer to the brightness levels of the left and right dots, with \(L\) indicating the low-brightness condition and \(H\) indicating the high-brightness condition. The additivity of the two factors can be inferred from MIC. However, MIC does not allow one to distinguish between different decisional stopping rules. Hence, we also used the SIC, which can be expressed as

\[
SIC = S_{(H,H)}(t) - S_{(H,L)}(t) - S_{(L,H)}(t) + S_{(L,L)}(t),
\]

where \(S(t)\) indicates the survivor function in one of the double-dot conditions. Combining MIC and SIC, one can make a strong inference regarding the processing architecture and the decisional stopping rule. Table 1 summarizes five possible models with different combinations of the processing architecture and the decisional stopping rule, as well as their MIC and SIC predictions (see Townsend & Nozawa, 1995).

Table 1 Summary of the five possible models and their MIC and SIC predictions

<table>
<thead>
<tr>
<th>Models</th>
<th>Processing Architecture</th>
<th>Stopping Rule</th>
<th>MIC</th>
<th>SIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial</td>
<td>Self-terminating</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Serial</td>
<td>Exhaustive</td>
<td>0</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Parallel</td>
<td>Self-terminating</td>
<td>&gt;0</td>
<td>&gt;0</td>
<td>-</td>
</tr>
<tr>
<td>Parallel</td>
<td>Exhaustive</td>
<td>&gt;0</td>
<td>&lt;0</td>
<td>+</td>
</tr>
</tbody>
</table>

\(-\rightarrow+\) indicates that the SIC goes from negative to positive values as a function of reaction time.
for $t > 0$, where $S_1(t)$, $S_2(t)$, and $S_{12}(t)$ represent the survivor function of the two single-dot and the double-dot conditions. If $C(t) < 1$, the capacity coefficient suggests limited-capacity processing. If $C(t) = 1$, the capacity coefficient suggests unlimited-capacity processing. If $C(t) > 1$, the capacity coefficient suggests supercapacity processing.

Moreover, two inequalities were computed to aid in the interpretation of the processing capacity. In both cases, we utilized a recent interpretation of these inequalities, developed by Townsend and Eidels (2011), in terms of the capacity coefficient. In that article, both inequalities for this task (and for the associated AND version of the task) were explicitly formalized within the same space as the capacity coefficient, allowing one to interpret the inequalities as bounds on unlimited-capacity performance.

The first inequality is the race-model inequality, which is also called the Miller inequality (Miller, 1982; Ulrich & Miller, 1997). The race-model inequality can be expressed as

$$ C(t) \leq \frac{\log[S_1(t) + S_2(t) - 1]}{\log[S_1(t) - S_2(t)]}. \quad (4) $$

The second inequality is the Grice inequality (Grice, Canham, & Boroughs, 1984; Grice, Canham, & Gwynne, 1984), which can be expressed as

$$ C(t) \geq \frac{\log[\text{MIN}\{S_1(t), S_2(t)\}]}{\log[S_1(t) - S_2(t)]}. \quad (5) $$

These inequalities place upper and lower boundaries on super- and limited-capacity processing, respectively. A violation of the race-model inequality indicates supercapacity processing, whereas a violation of the Grice inequality suggests very limited-capacity processing.

**General method**

**Participants**

Five undergraduate students (C.H., C.Y., L.R., Y.C., and Y.W.) from National Cheng Kung University participated in two experiments. A cue was manipulated to be either uninformative (50 % valid) or informative (100 % valid) to signal the onset of a target dot. The order of the two experiments was counterbalanced across the participants, with three of the participants (observers C.Y., Y.C., and Y.W.) first completing the 50 %-validity experiment, and the remaining (observers C.H. and L.R.) first completing the 100 %-validity experiment. Their ages ranged from 18 to 21 years old, with a mean age of 19. Each participant received NTD 1,200 for their participation. All participants had normal or corrected-to-normal vision.

**Apparatus**

The experiments were conducted in a darkened room with dim light. A personal computer with a 2.40-GHz Intel Pentium IV processor controlled the stimulus display and recorded the participants’ responses. The display resolution was 1,024 (pixel in width) × 768 (pixel in height). The visual stimuli were presented on a 19-in. CRT monitor (CTX VL951T) with a refresh rate of 85 Hz. E-Prime 1.1 was used to run the experiments (Schneider, Eschman, & Zuccolotto, 2002). The viewing distance was 60 cm. A chinrest was used in order to prevent head movements.

**Design, stimuli, and procedure**

At the beginning of each trial, a fixation point and two peripheral boxes (4° in height and width; the center-to-center distance between the two boxes was 12°) were presented for 500 ms (see Fig. 3 for an illustration of the experimental procedure). After the fixation point, a cue (a light gray square 3.2° in height and width) was presented at the center of one of the two peripheral boxes for 200 ms. Following a blank interval of 150 ms, a test display appeared for 100 ms, immediately followed by another blank interval of 1,900 ms. The participants had to quickly respond if they detected a target dot (the diameter of the dot stimulus was 0.2°); otherwise, no response was required. The intertrial interval was 1,000 ms.

We created four types of the test trials, including double-dot trials (both locations contain a target dot), two types of the single-dot trials (either the left or the right location contains a target dot), and no-dot trials (no dots are presented) (see Fig. 2). Each test trial type was equally probable. In addition, the brightness of the dot stimulus was manipulated in two levels, with equal presentation probabilities. Specifically, the luminance of the dot stimulus was either 0.3 cd/m² (high brightness) or 0.038 cd/m² (low brightness). The factorial combination of high and low brightness forms the HH, HL, LH, and LL conditions necessary for computation of the MIC and SIC.

In addition to the brightness manipulation, we manipulated the validity of an exogenous cue probing the target location. Since the cue was 100 % valid for the double-dot conditions and 100 % invalid for the no-dot condition, we manipulated the cue validity in the single-dot conditions.³ In the 50 %-validity experiment (uninformative cue), on half of the trials in

³ Due to the equal presentation frequencies of the four types of test trials (double-dot trials, two types of single-dot trials, and no-dot trials), the actual cue validities were 50 % in the 50 %-validity experiment (uninformative cue) and 75 % in the 100 %-validity experiment (informative cue).
the single-dot conditions a cue was followed by a dot stimulus, and on the remaining trials, a cue was followed by a blank. In the 100 %-validity experiment (informative cue), on all the trials in the single-dot conditions a cue was always followed by a target dot.

Each participant performed five sessions for each experiment, and each session was composed of ten blocks of 80 trials. The number of the test trials allowed us to collect more than 200 correct response times for each condition in order to estimate the RT distributions. Before each session began, each observer practiced a block of 80 trials. The length of the practice (approximate 5 min) was sufficient for dark adaptation.

Data analysis

The data analysis followed the suggestions of SFT (see Townsend & Nozawa, 1995). A critical assumption of SFT is that the brightness manipulation affects only the target location in which brightness was varied (i.e., varying the brightness of the left dot should not affect the RT for the right dot). This assumption is termed effective selective influence (Dzhafarov, 1999; Kujala & Dzhafarov, 2008; Townsend & Thomas, 1994).

Three levels of analyses were conducted to verify this assumption. First, mean RTs in the single-dot conditions were analyzed with $t$ tests. The selective-influence assumption holds when there are significant differences in the mean RTs between the HX and LX conditions and between the XH and XL conditions. Second, mean RTs in the double-dot conditions were analyzed with a $2 \times 2$ (brightness of the left dot) analysis of variance (ANOVA). If the selective-influence assumption were satisfied, we would expect to observe significant main effects of the brightness manipulation. Third, the RT distributions in the double-dot conditions were examined. If this assumption were satisfied, we would expect to observe that the survivor functions of the double-dot conditions were ordered at the factor levels.

Assuming that the selective-influence assumption is satisfied, we can further compute the MIC, SIC, and $C(t)$. We also conducted two analyses to confirm whether MIC was 0 (i.e., consistent with serial processing; see Table 1). First, the mean RTs in the double-dot conditions were analyzed with a two-way ANOVA. A significant interaction would suggest that MIC was not equal to 0. Second, a nonparametric bootstrapping method was used to simulate 1,000 samples for all of the double-dot conditions and to construct the 95 % confidence interval for MIC (see Van Zandt, 2000, for details). If the bootstrapped 95 % confidence interval for MIC included 0, the conclusion would be that MIC was 0. Moreover, we also utilized a nonparametric bootstrapping method to construct the 95 % confidence interval for the SICs, to determine whether the value of the SIC was 0 for all times $t$ (i.e., consistent with serial processing), and for the $C(t)$ function, to determine whether the value of $C(t)$ was equal to 1 for all times $t$ (i.e., consistent with unlimited-capacity processing).

Results

Accuracy was very high in both conditions, with the only nonzero error rates occurring in the no-dot trials. In the 50 %-validity condition, the highest error rate (i.e., false
alarm) was .02 (observer L.R.); in the 100 % validity condition, the highest error rate was .15 (observer L.C.). Although this error rate seems to indicate a possible response bias, we note that the next highest error rate was .04 (observer C.H.). Consequently, if there is a response bias, it was likely restricted to one observer, and we limited the remainder of our analyses to the RTs.

50 % cue validity

The data from the practice trials was excluded from analysis and correct RTs were extracted for further analysis. Table 2 shows the mean RTs in the double-dot conditions and MIC at the individual and group levels.

Prior to inferring the processing characteristics, we tested the selective-influence assumption. First, we found that the differences in mean RTs between the HX and LX conditions and the XH and XL conditions were all significant at both the individual and group levels (see Table 3 for results of the t tests). These results supported successful manipulations of brightness for the single-dot conditions. Second, mean RTs in the double-dot conditions were analyzed with a two-way ANOVA. The results showed that all the main effects of brightness manipulations were significant at the individual level and at the group level (see Table 4 for the ANOVA results). These results supported successful manipulations of brightness at the mean RT level. In addition, the interaction effects were also significant, thus suggesting nonadditive processing.

Finally, we tested the selective-influence assumption at the RT distribution level by observing the order of the survivor functions in the double-dot conditions at the factor levels. Figure 4 plots the survivor functions of the double-dot conditions at the individual and group levels. From visual inspection, we found that the survivor functions shifted from the HH, HL, and LH conditions to the LL condition. In addition, the survivor functions of the HH, HL, and LH conditions overlapped (primarily for observers C.Y. and L.R.). Table 5 presents the results of the Kolmogorov–Smirnov (K–S) tests on the survivor functions at the factor levels. The K–S results showed that the marginal distributions of the HH and HL conditions were significantly different from those of the LH and LL conditions, suggesting an effective manipulation of the brightness of the left dot. In addition, the marginal distributions of the HH and LH conditions were significantly different from those of the HL and LL conditions, suggesting an effective brightness manipulation of the right dot.

Taken together, the three levels of analyses converged to suggest that the selective-influence assumption held. Consequently, we then computed MIC and SIC at both the individual and group levels to infer the processing architecture and the decisional stopping rule.

All of the individual-level and group-level analyses showed similar patterns of results. The observed MIC was significantly greater than 0 (see Table 2). The results of ANOVA showed that the interaction effects were all significant (see Table 4). In addition, the nonparametric bootstrap results showed that the bootstrapped 95 % confidence intervals for MIC did not include 0 (see Fig. 5); thus, the conclusion is that MIC was greater than 0. These results suggest the existence of nonadditive or nonserial processing. Moreover, the nonparametric bootstrap results showed that the 95 % confidence intervals for SIC values were positive for all times t (see Fig. 6). When we combined the results from MIC and SIC, we inferred that all of the participants adopted parallel processing and followed a self-terminating stopping rule for decision making (see Table 1).

Next, we examined processing capacity (see Fig. 7). For most of the observers, the results showed that the bootstrapped 95 % confidence intervals for C(t) included 1 for the faster RTs but were below 1 for the slower RTs. These results suggest unlimited- to moderately limited-capacity processing. In these cases, however, the capacity was always above the lower Grice bound, indicating that capacity was in the range of what an unlimited-capacity parallel model would predict. Observer Y.C. was an exception (see Fig. 7d), and his results showed that the bootstrapped 95 % confidence interval for C(t) exceeded the race-model bound at the faster RTs, suggesting supercapacity processing.

In summary, the results of the 50 % validity experiment showed that MIC was greater than 0 and SIC values were positive for all times t, suggesting that all of the participants adopted parallel processing and followed a self-terminating stopping rule to detect the double dots. Although all of the participants showed the same processing architecture and decisional stopping rules, we found individual differences in their processing capacity. For observers C.H., C.Y., L.R., and Y.W., C(t) generally suggested unlimited- to moderately limited-capacity in information processing, which means that adding another signal might decrease or might not affect the efficiency of individual-channel processing. According to

<table>
<thead>
<tr>
<th>Observer</th>
<th>Condition</th>
<th>HH</th>
<th>HL</th>
<th>LH</th>
<th>LL</th>
<th>MIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.H.</td>
<td>224.64</td>
<td>228.29</td>
<td>226.55</td>
<td>238.65</td>
<td>8.45</td>
<td></td>
</tr>
<tr>
<td>C.Y.</td>
<td>301.79</td>
<td>301.66</td>
<td>301.98</td>
<td>323.50</td>
<td>21.65</td>
<td></td>
</tr>
<tr>
<td>L.R.</td>
<td>288.40</td>
<td>288.09</td>
<td>288.46</td>
<td>309.11</td>
<td>20.96</td>
<td></td>
</tr>
<tr>
<td>Y.C.</td>
<td>221.47</td>
<td>224.07</td>
<td>225.48</td>
<td>236.86</td>
<td>8.78</td>
<td></td>
</tr>
<tr>
<td>Y.W.</td>
<td>232.34</td>
<td>235.26</td>
<td>234.66</td>
<td>243.61</td>
<td>6.03</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>253.04</td>
<td>254.39</td>
<td>253.97</td>
<td>267.68</td>
<td>12.36</td>
<td></td>
</tr>
</tbody>
</table>
Eidels, Houpt, Altieri, Pei, and Townsend (2011), limited-capacity parallel processing suggests the potential existence of negative interactions at the stage of information accumulation. For observer Y.C., $C(t)$ suggested supercapacity to unlimited capacity in information processing, which means that adding another signal might increase or might not affect the efficiency of individual-channel processing.

100 % cue validity

We used the same criteria to analyze the data. Table 6 shows the mean RTs in the double-dot conditions and MIC at the individual and group levels.

We first tested the selective-influence assumption. The differences in mean RTs between the HX and LX conditions and between the XH and XL conditions were significant at both the individual level and the group level (see Table 7 for results of the $t$ tests). These results supported successful manipulations of brightness for the single-dot conditions. Second, the mean RTs in the double-dot conditions were analyzed with a two-way ANOVA. The results showed that all of the main effects of brightness manipulations were significant (see Table 8 for the ANOVA results). These results supported successful manipulations of brightness at the mean-RT level. Unlike the 50 %-validity experiment, however, the interaction effects were not significant (see Table 8), suggesting additive or serial processing.

Finally, we tested the selective-influence assumption at the RT distribution level. Figure 8 plots the survivor functions of the double-dot conditions at both the individual and group levels. From visual inspection, we found that the survivor functions gradually shifted from the HH, HL, and LH conditions to the LL condition. In addition, the survivor functions of the HL and LH conditions overlapped. Table 9 presents the results of the K–S tests on the survivor functions at the factor levels. The K–S results showed significant differences between the marginal distributions of the HH and HL conditions and those of the LH and LL conditions, and between the marginal distributions of the HH and LH conditions and those of the HL and LL conditions. These results supported effective brightness manipulations for the left and right dots.

Because the selective-influence assumption held, we computed MIC and SIC at both the individual and group levels. All of the individual-level and group-level analyses showed similar patterns of results. The observed MIC was about 0 (see Table 6), and the results of the ANOVA showed that the interaction effects were not significant (see Table 8). In addition, the nonparametric bootstrap results showed that the bootstrapped 95 % confidence intervals for MIC included 0 (see Fig. 9). These results suggest the existence of additive or serial processing. Moreover, the nonparametric bootstrap results showed that the 95 % confidence intervals for SIC values were 0 for all times $t$ (see Fig. 10). When we combined the results from MIC and SIC, we inferred that all of the

---

**Table 3** Results of the $t$ tests on the mean reaction times of the single-dot conditions in the 50 %-validity experiment

<table>
<thead>
<tr>
<th>Observer</th>
<th>Mean RT (ms)</th>
<th>HX vs. LX</th>
<th>XH vs. XL</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.H.</td>
<td>10.52***</td>
<td>3.40***</td>
<td>3.03**</td>
</tr>
<tr>
<td>C.Y.</td>
<td>498</td>
<td>499</td>
<td>497</td>
</tr>
<tr>
<td>L.R.</td>
<td>8.83***</td>
<td>4.31***</td>
<td>2.99**</td>
</tr>
<tr>
<td>Y.C.</td>
<td>499</td>
<td>499</td>
<td>498</td>
</tr>
<tr>
<td>Y.W.</td>
<td>2.64**</td>
<td>9.89***</td>
<td>10.57***</td>
</tr>
<tr>
<td>Group</td>
<td>10.92***</td>
<td>2.64**</td>
<td>11.95***</td>
</tr>
</tbody>
</table>

$t$ represents the $t$ statistics; $df$ represents the degrees of freedom. $^*$ $p < .05$. $^{**}$ $p < .01$. $^{***}$ $p < .001$.

**Table 4** Results of the ANOVA in the 50 %-validity experiment

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Right $F$</td>
<td>12.92***</td>
<td>4.67*</td>
<td>4.90*</td>
<td>15.18***</td>
<td>19.38***</td>
<td>20.55***</td>
<td></td>
</tr>
<tr>
<td>Left $F$</td>
<td>21.29***</td>
<td>4.95*</td>
<td>4.55*</td>
<td>21.96***</td>
<td>15.66***</td>
<td>23.18***</td>
<td></td>
</tr>
<tr>
<td>MSE $F$</td>
<td>6.13***</td>
<td>4.78*</td>
<td>4.84*</td>
<td>5.99*</td>
<td>5.01*</td>
<td>13.83***</td>
<td></td>
</tr>
<tr>
<td>MSE $1267.35$</td>
<td>5.667.81</td>
<td>803.84</td>
<td>454.19</td>
<td>3937.27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$df$</td>
<td>994</td>
<td>996</td>
<td>995</td>
<td>996</td>
<td>996</td>
<td>4,993</td>
<td></td>
</tr>
</tbody>
</table>

$F$ represents the $F$ statistics; MSE represents the mean squared error; $df$ represents the degrees of freedom of the denominator; the degree of freedom of the numerator was 1. $^*$ $p < .05$. $^{**}$ $p < .01$. $^{***}$ $p < .001$.
participants adopted serial processing and followed a self-terminating stopping rule for decision making (see Table 1).

Figure 11 shows the results for processing capacity. The results from all observers except observer Y.W. and the group data (see Fig. 11a, b, c, d, and f) showed that the bootstrapped 95% confidence intervals for $C(t)$ were below 1 for all times $t$. The bootstrapped 95% confidence interval did not exceed the race-model bound and the Grice bound. These results suggested that most of the participants detected double dots with moderately limited-capacity processing. Observer Y.W. was an exception. The bootstrapped 95% confidence interval for $C(t)$ included 1 for the faster RTs, but below 1 for the slower RTs, suggesting unlimited- to limited-capacity processing (see Fig. 11e).

It is worthwhile to note that the standard serial model (assuming independent and identically distributed processing times for each stage of the process) predicts that $C(t)$ should be equal to 1/2 (Townsend & Nozawa, 1995). Likewise, if participants utilized the cue in an unbiased fashion (i.e., always attended to the cued location), then across trials, no bias should occur toward either of the locations, and the Grice

![Fig. 4](image)

**Fig. 4** Plots of the survivor functions of the double-dot conditions in the 50%-validity experiment. HH represents that two dots were both in high brightness (solid line); HL represents that the left dot was in high brightness and the right dot was in low brightness (dotted line); LH represents that the left dot was in low brightness and the right dot was in high brightness (dotted line); and LL represents that the two dots were both in low brightness (dash-dotted line).

![Table 5](image)

**Table 5** Results of the Kolmogorov–Smirnov tests on the survivor functions of the double-dot conditions at the factor levels in the 50%-validity experiment.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>1.43*</td>
<td>1.55*</td>
<td>1.47*</td>
<td>1.89**</td>
<td>2.01**</td>
<td>2.54***</td>
</tr>
<tr>
<td>Right</td>
<td>1.75**</td>
<td>1.45*</td>
<td>1.42*</td>
<td>1.77**</td>
<td>1.39*</td>
<td>2.62***</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001
bound should also be equal to 1/2. Examination of Fig. 11 reveals that for most of our observers (four of five), $C(t)$ and the Grice bound are both approximately equal to 1/2, providing quite strong evidence of serial processing in the 100 %-validity experiment.

In summary, the results of the 100 %-validity experiment showed that for all participants, MIC was equal to 0 (or, more correctly, could not be significantly distinguished from 0), and the SIC values were equal to 0 for all times $t$, suggesting that the participants adopted serial processing and followed a self-terminating stopping rule to detect the double dots. In the present case, rather than reflecting an ordered processing of possible locations, a serial self-terminating processing strategy reflects the exclusive processing of a single location, in line with the idea that attention is narrowly focused only on the cued location. $C(t)$ generally suggests limited capacity in information processing, which means that adding another signal decreased the efficiency of processing on the original signal. Observer Y.W. showed unlimited to limited capacity. Serial processing with unlimited capacity suggests that the average individual processing time is constant, regardless of the number of targets to be processed (Townsend, 1990; Townsend & Ashby, 1983).

General discussion

We conducted two experiments to examine whether the manipulation of cue validity would affect the decision process of detecting the redundant targets. A combination of a redundant-target detection task and Posner’s spatial-cueing paradigm was used, with cue validity being manipulated between experiments. We followed the suggestions of SFT to infer the processing architecture, the decisional stopping rule, and the processing capacity of
target detection. To summarize the results, the processing architecture and processing capacity were affected by the validity of a location cue, but the decisional stopping rule was not. In the 50 %-validity experiment, a parallel self-terminating processing was adopted. By contrast, serial self-terminating processing was adopted in the 100 %-validity experiment. In addition, processing capacity was generally more limited in the 100 %-validity experiment than in the 50 %-validity experiment, reflecting a switch from parallel to serial processing (see Townsend & Nozawa, 1995).

The present results go beyond those of Yang et al. (2013); in that study, they varied the relative frequency of feature changes in a change detection task and found that observers preferentially processed a feature with a higher change probability in a serial fashion. Here, we varied the validity of a location cue; this manipulation allowed us to link the present results to the previous literature on models of cued attention and attentional focus. Taken together, these results support the relative saliency hypothesis (Yang, 2011; Yang et al., 2013; Yang et al., 2011) and suggest a fundamental link between attention to a cued location (i.e., the type of “space”-based attention normally discussed in the visual-attention literature) and attentional weighting of stimulus features (i.e., the type of “feature”-based attention normally discussed in conjunction with higher-level cognitive tasks such as categorization). We will discuss each of these in turn.

Top-down versus bottom-up attentional control

The present results address the question of whether top-down attention to valid cues (such as the validity of each of the location cues) can override bottom-up attention to salient cues (such as the onset of the cue regardless of its validity).
Although this issue has been debated for a few years (Folk & Remington, 2010; Folk, Remington, & Johnston, 1992; Theeuwes, 1991, 1994, 2004b, for different views), less is known about how top-down attention and bottom-up attention interactively affect the time course of the decision process during target detection. Our results show that the information-processing strategy varied according to the cue validity. These findings ruled out the possibility that attention to any onset cue is mandatory (Schreij et al., 2008; Theeuwes, 2004b). If this were the case, we would have expected processing to be serial in both the 50 %- and 100 %-validity experiments. In addition, the results also ruled out the possibility that a redundancy gain offers an advantage for parallel processing.

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Mean reaction times (in milliseconds) of the double-dot conditions and MIC in the 100 %-validity experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer</td>
<td>Condition</td>
</tr>
<tr>
<td></td>
<td>HH</td>
</tr>
<tr>
<td>C.H.</td>
<td>230.24</td>
</tr>
<tr>
<td>C.Y.</td>
<td>276.68</td>
</tr>
<tr>
<td>L.R.</td>
<td>257.31</td>
</tr>
<tr>
<td>Y.C.</td>
<td>219.20</td>
</tr>
<tr>
<td>Y.W.</td>
<td>233.07</td>
</tr>
<tr>
<td>Group</td>
<td>244.15</td>
</tr>
</tbody>
</table>

Fig. 7 Plots of the capacity coefficient in the 50 %-validity experiment. The thick solid lines represent $C(t)$; the thin solid lines represent the 95 % confidence interval for $C(t)$; the dashed lines represent the Grice bound; and the dash-dotted lines represent the race-model bound. For all participants except observer Y.C., the 95 % confidence intervals for $C(t)$ included 1 for the faster reaction times (RTs) but were below 1 for the slower RTs, suggesting unlimited to moderately limited-capacity processing. For observer Y.C. (see Fig. 6d), his results showed that the bootstrapped 95 % confidence interval for $C(t)$ exceeded the race-model bound at the faster RTs, suggesting supercapacity processing.
processing of both locations over serial location processing. If this were the case, participants would have strategically ignored the location cues, and we would have expected processing to be parallel, regardless of the manipulation of cue validity.

Our results support a model of attentional focus in which participants can exert full top-down control over where to attend. When a location cue was uninformative (50% valid), participants did not utilize the cue to detect the targets. To optimize their detection performance, participants strategically ignored the cue and adopted a strategy that focused on all of the items in the display (i.e., a wide attentional focus), resulting in a parallel processing strategy. In contrast, when a location cue was informative (100% valid), the participants only paid attention to the cued location to detect the target (i.e., a narrow attentional window), resulting in exclusive processing of the cued location. A likely explanation for the results of the two experiments is that the cue was highly predictable on each trial, such that effective top-down control could be executed over the selection of the decision strategy. For instance, Theeuwes and Burger (1998) found that top-down attention could overcome an irrelevant distractor so long as the distractor was fixed and of a known value. This explains why an uninformative cue could be ignored in the present task and implicates top-down control over attention to the informative cue, resulting in a serial decision strategy.

Table 7 Results of the t tests on the mean reaction times of the single-dot conditions in the 100 %-validity experiment

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HX vs. LX t</td>
<td>4.80***</td>
<td>6.30***</td>
<td>6.49***</td>
<td>6.81***</td>
<td>7.99***</td>
<td>13.79***</td>
</tr>
<tr>
<td>df</td>
<td>499</td>
<td>497</td>
<td>498</td>
<td>497</td>
<td>497</td>
<td>2,494</td>
</tr>
<tr>
<td>XH vs. XL t</td>
<td>4.41***</td>
<td>5.93***</td>
<td>8.87***</td>
<td>8.48***</td>
<td>11.38***</td>
<td>15.36***</td>
</tr>
<tr>
<td>df</td>
<td>499</td>
<td>498</td>
<td>498</td>
<td>499</td>
<td>495</td>
<td>2,493</td>
</tr>
</tbody>
</table>

Table 8 Results of the ANOVA in the 100 %-validity experiment

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Left     F</td>
<td>4.07*</td>
<td>10.18**</td>
<td>10.65**</td>
<td>9.33**</td>
<td>12.19***</td>
<td>31.48***</td>
</tr>
<tr>
<td>df</td>
<td>996</td>
<td>995</td>
<td>995</td>
<td>996</td>
<td>991</td>
<td>4989</td>
</tr>
<tr>
<td>Right    F</td>
<td>4.59*</td>
<td>4.49*</td>
<td>6.29*</td>
<td>10.45**</td>
<td>25.94***</td>
<td>25.98***</td>
</tr>
<tr>
<td>df</td>
<td>995</td>
<td>995</td>
<td>995</td>
<td>995</td>
<td>995</td>
<td>995</td>
</tr>
<tr>
<td>Left * Right F</td>
<td>0.28</td>
<td>0.82</td>
<td>0.25</td>
<td>1.65</td>
<td>0.31</td>
<td>0.09</td>
</tr>
<tr>
<td>MSE</td>
<td>1,485.62</td>
<td>2,891.60</td>
<td>3,455.68</td>
<td>1,188.19</td>
<td>436.27</td>
<td>2,426.12</td>
</tr>
<tr>
<td>df_e</td>
<td>996</td>
<td>995</td>
<td>995</td>
<td>996</td>
<td>991</td>
<td>4989</td>
</tr>
</tbody>
</table>

F represents the F statistics; MSE represents the mean squared error; df_e represents the degrees of freedom of the denominator; the degree of freedom of the numerator was 1. * p < .05. ** p < .01. *** p < .001

Implication for theories of selective attention in detection

It is useful to consider the implications of the present results for models of attentional focus in decision making. Because the modeling implications are based on nonparametric analyses, they place constraints on the types of parametric models that might be designed to account for the focusing of attention. For instance, processing in the 50% validity experiment was found to be consistent with a parallel-processing model and to effectively rule out a serial-processing model, and also a coactive-processing model. This implies that parametric models that assume either serial processing of locations or, more plausibly, a pooling of information from both locations into a single channel are likely incorrect in those assumptions. This speaks against several models that assume a single decision mechanism, such as a simple signal detection model that combines all of the information into a single decision (Green & Swets, 1966).

To take a further example, Eckstein and colleagues (Eckstein, Pham, & Shimozaki, 2004; Eckstein, Shimozaki, & Abbey, 2002) proposed a Bayesian observer model of cued detection, in which each location is weighted by the prior probability of the target appearing at each location. Crucially, according to this theory, a likelihood ratio of both locations is used to drive a single decision mechanism (i.e., comparing the ratio to some threshold). The authors of this model did not develop their model to predict RTs,
aiming instead to capture patterns of hit and false alarm rates, but if one were to adopt a simple assumption that the likelihood ratio drives a single-channel sequential-sampling model (e.g., in the manner of the exemplar-based random-walk model; Nosofsky & Palmeri, 1997), then this model would predict that processing of both locations would occur in a coactive fashion. Here we found no evidence of coactivity in either experiment, and furthermore, Fifić, Little, and Nosofsky (2010) showed that even an extremely flexible single-channel model, which freely estimates the sampling rate for each of the factorially combined items, cannot mimic a serial- or a parallel-processing model, ruling out a large class of models based on pooling into a single decision channel.

Implication for theories of selective attention in other domains

The shift from diffuse, parallel attentional processing when the cue was uninformative (comparable to when there is no cue at all) to focused, serial processing when the cue was informative is reminiscent of the validity effects that have

![Fig. 8](image-url)  
**Fig. 8** Plots of the survivor functions of the double-dot conditions in the 100 %-validity experiment. HH represents that two dots were both in high brightness (solid line); HL represents that the left dot was in high brightness and the right dot was in low brightness (dotted line); LH represents that the left dot was in low brightness and the right dot was in high brightness (dashed line); and LL represents that two dots were both in low brightness (dash-dotted line).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>1.38*</td>
<td>1.48*</td>
<td>1.97**</td>
<td>1.55*</td>
<td>2.08**</td>
<td>2.76***</td>
</tr>
<tr>
<td>right</td>
<td>1.67**</td>
<td>1.83**</td>
<td>1.44*</td>
<td>1.70**</td>
<td>1.43*</td>
<td>2.47***</td>
</tr>
</tbody>
</table>

*p < .05. *p < .01. ***p < .001

Table 9 Results of the Kolmogorov–Smirnov tests on the survivor functions of the double-dot conditions at the factor levels in the 100 %-validity experiment
been found in other psychological tasks. The ability to preferentially weight important or salient stimulus attributes would allow people to efficiently and optimally combine information when making decisions. The trade-off between salience—how much a stimulus attribute grabs your attention—and validity—how important a stimulus attribute is to achieving some goal (e.g., making a correct decision)—has been extensively studied in categorization and multiple-cue probability learning (for reviews, see Kruschke & Johansen, 1999; Little & Lewandowsky, 2012). The most important findings of that literature are that salient dimensions and valid dimensions are weighted more than are nonsalient or nonvalid dimensions in decisions. Furthermore, these two factors trade off in an additive fashion (Kruschke & Johansen, 1999).

These factors also play roles in other tasks, such as feature matching (Yang, Chiu, & Yeh, 2012), visual search, and change detection (Theeuwes, 2004a, 2010; Woodman & Luck, 2007; Yang, 2011; Yang et al., 2011), and are reminiscent of the Bayesian model of attentional focus described above (Eckstein et al., 2004; Eckstein et al., 2002).

In categorization, several prominent theories assume that selective attention changes the similarity relations between stimuli, but that all stimulus information is integrated and used to drive decision making (e.g., Kruschke, 1996; Nosofsky, 1986; Nosofsky & Palmeri, 1997). This weighting is typically instantiated as a multiplier on a dimension that is higher for more-valid dimensions (e.g., as in the generalized context model, or GCM; Nosofsky, 1986) or more-salient dimensions (Kruschke & Johansen, 1999). Taken together with models of RTs in categorization, attended dimensions exert more weight in driving the decision, but this occurs through a process of “stretching” the attended dimension. All of the stimulus dimensions are still incorporated into the decision, which is driven by the summed similarity of a probe stimulus to the previously seen stimuli stored in memory. Decision making, in these
theories, occurs in a pooled or coactive fashion (e.g., Fifić et al., 2010; Fifić, Nosofsky, & Townsend, 2008; Little, Nosofsky, Donkin, & Denton, 2013).

On the other hand, studies requiring fast decisions under time pressure have shown that only salient information is available early in the decision process, suggesting that complete integration of information is not possible under some conditions, at least for separable stimulus attributes (Cohen & Nosofsky, 2003; Lamberts, 1995, 1998, 2000). In a series of studies, Lamberts (1995, 1998, 2000) manipulated the amount of time available for categorizing stimuli that had both salient dimensions and valid dimensions (i.e., in these studies, the valid dimensions were defined by virtue of an item’s position in the category space, but salient dimensions were those that attracted attention, regardless of their validity). Under short deadline or short response signals, Lamberts found that the decisions were driven primarily by the salient stimulus information. Only at longer response deadlines did valid stimulus information exert an influence on responding.

Lamberts’s (1995, 1998, 2000) results (recently reviewed by Kent, Guest, Adelman, & Lamberts, 2014) provide a link between studies of categorization and attentional focusing, as we studied here. The initial processing of salient information could arise due to preferential serial processing of that information. Such an account is akin to an orienting theory of attention (e.g., Sewell & Smith, 2012) that proceeds serially from salient to valid information; under time-constrained response conditions, only the first-processed salient information is available for decision making. On the other hand, an alternative possibility is that salient and valid dimensions are processed in parallel, but that salient information is processed faster than valid information.

To differentiate the two accounts, we can draw on recent evidence showing that for spatially separated stimulus attributes such as those used in Lamberts’s studies,
dimensional processing proceeds in a serial and self-terminating fashion (Fišić et al., 2010; Little, Nosofsky, & Denton, 2011). Taken together, this would implicate a serial selective-attention mechanism: Salient stimulus information is processed before valid stimulus information. The implication is that for information of equal salience in which the salient cues are highly predictable (Theeuwes & Burger, 1998), attention to stimulus dimensions of differing validities would also proceed in a serial fashion. Consequently, cues of high validity should attract attention more than would cues of low validity, and on the basis of the present results, we would expect that attentional allocation would change the nature of information processing. This conceptualization of selective attention is attractive because it coheres with eyetracking studies that have shown that fixation is correlated with attention weights in categorization models like the GCM (Blair, Watson, Walshe, & Maj, 2009; Hoffman & Rehder, 2010; Rehder & Hoffman, 2005a, 2005b) and explains why attention to one stimulus attribute can result in total ignorance of an alternative stimulus attribute. For instance, participants fail to notice nonrelevant correlations in their environment when other, relevant cues can be perfectly utilized to make decisions (Little & Lewandowsky, 2009).
Future directions

In this article, we have identified a strategic trade-off between the focused attention, serial processing in the 100 %-validity experiment and diffuse attention, parallel processing in the 50 %-validity experiment. What this suggests is that observers are sensitive to the relative benefits of processing a single target, the redundancy gains that can be achieved when there are two targets and when to switch from one strategy to the other. For example, LaBerge (1983) showed that attentional focus can vary in size depending on the extent of the task-relevant stimulus (e.g., the width of one letter in a letter identification task or the width of a five-letter word in a word categorization task). However, the larger the spatial extent of the attentional focus the greater is the reduction in the overall level of neural activation (Muller, Bartelt, Donner, Villringer, & Brandt, 2003). This finding is coherent with models that assume that the total number of attentional resources is limited and vary by the size of the attentional focus (see Sewell & Smith, 2012, for a review). This suggests that if the benefits of processing a single target do not outweigh the minimum time benefits of parallel processing of both locations, then observers are unlikely to switch from parallel to serial processing.

One manipulation that could test this prediction would be to simply decrease the spatial extent of the double target such that both targets can be processed within a smaller attentional window. Of course, there are likely lower limits on how close two targets can be before they are no longer processed in an independent fashion. An alternative suggestion is to utilize the same double-dot detection paradigm but require that both targets be present before a “yes” response is emitted (i.e., an AND decision rule). Under those conditions, processing of both target locations is mandatory and a self-terminating rule cannot be applied. Consequently, the parallel processing redundancy benefits are likely to outweigh any benefits of spatial cueing. Future experiments should examine this prediction.

Conclusion

In conclusion, the present study suggests that spatial cueing influences the processing architecture and the processing capacity of the decision process when detecting redundant targets. More importantly, this study has demonstrated the flexibility of the decision mechanism in cued detection, supporting the idea of top-down attentional control, along with attentional focusing on the detection decision process.

Author Note We thank T.-C. Wu for his assistance on figure modification. This work was supported by grants from National Science Council to C.-T.Y. (NSC 100-2410-H-006-028-MY2 and NSC 102-2628-H-006-001-MY3) and NCKU top-notch project proposal to C.-T.Y. Parts of the results were presented at the 45th Annual Meeting of the Society for Mathematical Psychology in Columbus, Ohio.

References


