Dividing attention between two transparent motion surfaces results in a failure of selective attention

Zachary Raymond Ernst
Department of Psychology, University of Washington, Seattle, Washington, USA

John Palmer
Department of Psychology, University of Washington, Seattle, Washington, USA

Geoffrey M. Boynton
Department of Psychology, University of Washington, Seattle, Washington, USA

In object-based attention, it is easier to divide attention between features within a single object than between features across objects. In this study we test the prediction of several capacity models in order to best characterize the cost to dividing attention between objects. Here we studied behavioral performance on a divided attention task in which subjects attended to the motion and luminance of overlapping random dot kinematograms, specifically red upward moving dots superimposed with green downward moving dots. Subjects were required to detect brief changes (transients) in the motion or luminance within the same surface or across different surfaces. There were two primary results. First, the dual-task deficit was large when attention was divided across two surfaces and near zero when attention was divided within a surface. This is consistent with limited-capacity processing across surfaces and unlimited-capacity processing within a surface—a pattern predicted by established theories of object-based attention. Second and unexpectedly, there was evidence of crosstalk between features: when cued to monitor transients on one surface, response rates were inflated by the presence of a transient on the other surface. Such crosstalk is a failure of selective attention between surfaces.

Keywords: transparent motion, divided attention, object-based attention

Citation: Ernst, Z. R., Palmer, J., & Boynton, G. M. (2012). Dividing attention between two transparent motion surfaces results in a failure of selective attention. *Journal of Vision*, *12*(12):6, 1–17, http://www.journalofvision.org/content/12/12/6, doi:10.1167/12.12.6.

Introduction

More sensory information is available to the visual system than can be effectively processed. Due to these limits in processing, information competes for memory encoding, perceptual decisions, and motor responses. Visual attention helps resolve this competition by selecting relevant information on the basis of spatial location, feature, or object, which biases the sensory processing in favor of the behaviorally relevant information (Desimone, 1998; Desimone & Duncan, 1995; Kanwisher & Wojciulik, 2000; Posner, 1980). Object-based attention is hypothesized to select all the features of a behaviorally relevant object, serving to improve the encoding of its component features relative to an unattended object (Duncan, 1984; Kahneman, Treisman, & Gibbs, 1992; Treisman, 1998; Valdes-Sosa, Cobo, & Pinilla, 1998). When behaviorally relevant features belong to different objects, can observers selectively attend to those features, or does object-based attention interfere with the selection process?

Duncan (1984) proposed that selective attention operates at the object-based level, limiting selection to one object at a time. Using a dual-task paradigm, subjects performed multiple perceptual judgments regarding features within one object or between two objects. Performance was better when the judgments regarded features belonging to the same object versus different objects (Duncan, 1984). Vecera and Farah (1994) later demonstrated similar divided attention costs—using Duncan's stimuli—regardless of whether the two objects were superimposed or separated in space, lending further support that object ownership limits perception. Ideally spatial separation could be better controlled for within the stimulus. In addition, the types of feature judgments made within and between objects were not the same, leading to the concern that task demands may have been different across conditions.

Valdes-Sosa, Cobo, and Pinilla (2000) improved the paradigm for studying object-based attention by

doi: 10.1167/12.12.6 Received May 4, 2012; published November 12, 2012 ISSN 1534-7362 © 2012 ARVO

superimposing two random dot fields that rotated in opposite directions. As is typical with a random dot kinematogram, each dot existed for a limited number of frames before being randomly redrawn, forcing surface segregation to rely on the global motion of the dots. Surface selection is therefore exclusively feature-based since dots belonging to either field were spatially intermingled. Subjects reported the direction of two brief translational probes, which occurred in either surface and with a variable interprobe interval. Performance in reporting the direction of the second probe dropped when it occurred in the other surface within 600 ms or less of the first. The authors described the first probe as an exogenous cue, which captured object-based attention decreasing the subjects' sensitivity to probes on the second surface. These results provide evidence that perception of objects is capacity limited. The authors describe the capacity limit as a "difficulty in switching attention rapidly between surfaces" (Valdes-Sosa et al., 2000).

Multiple groups have argued that capacity is unlimited when dividing attention within an object, but requires switching when attempting to divide attention between objects (Blaser, Pylyshyn, & Holcombe, 2000; Duncan, 1984; Valdes-Sosa et al., 2000). In a dual-task paradigm where observers are asked to divided their attention across multiple objects, an all-or-none switching model assumes that on a given trial the observer is constrained to select only one object at a time (Bonnel & Haftser, 1998; Sperling & Melchner, 1978) and therefore must guess when asked to recall the properties of a second object. Thus the all-or-none switching model predicts a negative trial-by-trial covariance (Bonnel & Prinzmetal, 1998) and a decrease in overall performance known as a dual-task cost. In addition to the all-or-none switching model, there are limited-capacity parallel models that also predict a dual-task cost, but zero trial-by-trial covariance. One specific limited-capacity model is the fixed-capacity model, which maintains a fixed amount of information processing when attention is divided (Shaw, 1980). These models make specific predictions of dual-task performance from baseline single-task performance. In this paper we measured dual-task performance when attention was divided within versus between two surfaces and compared behavioral performance to the predictions of two common capacity models at either end of the capacity continuum: the unlimited-capacity parallel and the all-or-none switching models.

Most capacity models assume perfect selection, but a dual-task deficit could also arise from selective attention errors cause by distractor interference. The difficulty in dividing attention between objects may arise from interaction between feature channels through *crosstalk* (Navon & Miller, 1987). If object-based attention facilitates the selection of all of an object's features, then task-irrelevant features may interfere when attempting to select specific features from multiple objects (Davis, Driver, Pavani, & Shepherd, 2000). The interference due to crosstalk may increase when attention is divided between objects composed of competing features within a feature dimension (e.g., different directions of motion). The prevalence of selection errors observed in our data leads us to propose a new capacity model that takes into account crosstalk.

In order to measure the capacity of divided attention within and between objects, we measured accuracy when attention was divided across features within and between two superimposed transparent motion surfaces created from a random dot kinematogram. Following in the tradition of Duncan and Valdes-Sosa, we chose to focus on accuracy rather than reaction times in order to test the predictions made by specific capacity models. Although our model does not make specific reaction time predictions, dual-task deficits may also manifest in slower reaction times when attention is divided between surfaces (Lamy & Egeth, 2002; Watson & Kramer, 1999). In addition, interference due to crosstalk has also been shown to affect reaction times (Navon & Miller, 1987; Treisman, Kahneman, & Burkell, 1983).

Methods

Participants

Five subjects participated in this study, including the first author. All subjects gave written informed consent in accord with the human subject protocol at the University of Washington (Seattle, WA).

Apparatus

Stimuli were presented on a CRT monitor with a resolution of 1024 × 768 pixels, and viewed from a distance of 57 cm. Subject responses were collected by keyboard button presses. The code was written in MATLAB (MathWorks, Natick, MA) and presented using Psychtoolbox (http://www.psychtoolbox.org) (Brainard, 1997; Pelli, 1997) on a computer running Windows 7.

Stimuli

The stimuli consisted of two superimposed moving surfaces composed of randomly drawn dots with a unique color-motion conjunction, i.e., red-up and green-down or vice versa, counter-balanced across

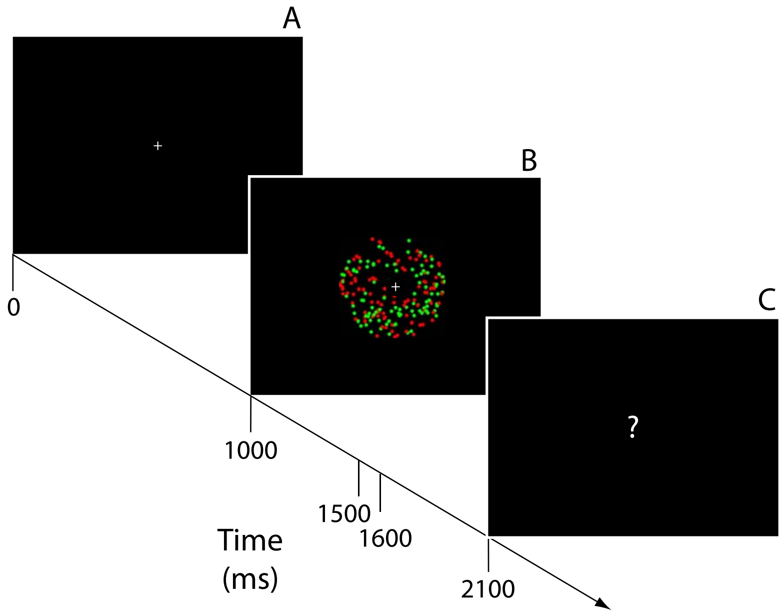

Figure 1. Trial structure. (A) Each trial commenced with a 1000 ms pretrial interval with a fixation cross. (B) The stimulus then appeared for 1100 ms. At 500 ms, poststimulus onset, 0–4 100 ms luminance/motion transients could occur. (C) Following the stimulus, the observer made one or two yes/no responses without time pressure.

sessions. Each surface was composed of 100 dots. To remove a potential depth cue, the depth order of overlapping dots (which dots occludes the other dot) was randomized. The diameter of the dots was 0.8°. The dots were confined to annulus with an inner diameter of 3° and an outer diameter of 16° (see Figure 1B). A fixation plus was placed at the center of the annulus. The dots moved coherently at a rate of 8°/s. Each dot was presented with a limited lifetime of 12 frames (200 ms), and was subsequently redrawn at a random position. The luminance of the green dots was reduced to match the luminance of the red dots at the maximum intensity of the red channel (33 cd/m^2). The (x, y) CIE 1931 xyz space coordinates for the red and green dot colors were (0.612, 0.331) and (0.279, 0.581), respectively. The monitor background was set to black with a luminance of less than 1 cd/m^2.

Procedure

Prior to a block of trials, subjects received specific instructions regarding which feature, or pair of features, to attend in order to perform one or two detection tasks. There were a total of four conditions: two single-task and two dual-task conditions. For the *single-task motion* condition, subjects were cued to attend to the speed of one of the two surfaces (e.g., "attend to the speed of the upward-moving red surface"). For the *single-task luminance* condition, subjects were cued to attend to the luminance of one of the two surfaces (e.g., "attend to the brightness of the red surface"). For the *dual-task, within-surface* condition, subjects were cued to attend to the speed and the luminance of one of the two surfaces (e.g., "attend to the speed AND brightness of the upward-moving red surface"). For the *dual-task, between-surface* condition, subjects were cued to attend to the speed of one surface and the luminance of the second surface (e.g., "attend to the speed of the upward-moving surface AND the brightness of the green surface").

The trial structure is schematized in Figure 1. Each trial began with a 1000 ms pretrial interval, consisting of a fixation plus centered on a black screen (Figure 1A). The stimulus movie followed for 1100 ms (Figure 1B). Stimulus "transients" occurred 500 ms after the onset of the moving surfaces and consisted of brief (100 ms) decrements in speed and/or luminance of the dots within each surface (Figure 1B). On every trial, there was an independent 50% chance of a transient occurring in each of the four features. Thus, from zero to four stimulus transients occurred on each trial.

Following stimulus offset, subjects were queried to indicate whether or not a transient occurred in the cued feature(s) by pressing one of two buttons (a yes/no response). There was no time pressure. To help reduce response errors, the yes/no responses for the motion and luminance tasks were mapped to separate pairs of keyboard buttons; subjects used their left and right pointer fingers to perform the motion and luminance tasks (e.g., "did the speed of the upward-moving surface decrease, yes (z) or no (x)?" and "did the brightness of the red surface decrease, yes (.) or no (/)". The order of report, whether motion or luminance was

	Speed decrement (%)		Luminance decrement (%)	
Subject	Upward	Downward	Red field	Green field
1	43	43	36	36
2	36	36	26	29
3	63	50	34	38
4	45	45	33	36
5	26	26	27	36

Table 1. Percent of speed or luminance decrement from baseline intensity (8°/sec and 33 cd/m^2) for transient events. The percentages are tabulated for each surface feature within each subject.

queried first, was pseudo-randomized across trials to prevent response biases.

Subjects practiced all four experimental conditions over the course of two to three one-hour sessions until they reported feeling comfortable with the task. Experimental data was then collected over two one-hour sessions. The magnitudes of the speed and luminance decrements were adjusted during the training sessions to ensure single-task performance levels above 80% correct but below ceiling. Intensities decrements were chosen separately for each surface (Table 1). For the motion task, Subject 3 exhibited much higher sensitivity for upward motion decrements then downward motion decrements. For the luminance task, four subjects (2–5) exhibited slightly higher sensitivity for the luminance decrements in the green field than in the red field.

During the experimental sessions, the magnitude of the speed and luminance decrements was held constant. Subjects performed blocks of 32 trials, preceded by specific attention instructions. Each attention condition was grouped into sets of four blocks. For example, for the single-task motion, subjects alternated between attending upward motion for 32 trials and downward motion for 32 trials. After four blocks (128 trials total), the subject began a new (randomly selected) cue condition. In this manner, 256 trials were collected for each of the four cue conditions in each one-hour session for a total of 512 trials per cue condition.

Analysis

We collapsed performance across surfaces because we were not interested in performance differences within a feature dimension (e.g., upward vs. downward motion or red vs. green). Behavioral performance was analyzed at several different levels. For a coarse analysis of performance, we averaged across the hit rate and correct rejection rate to compute a percent correct for each condition. Going further, we analyzed the joint dual-task performance for signs of independence between tasks. At the finest level of analysis, we compared hit rates to false alarm rates conditionalized on each type of stimulus transient. Finally, we fit a parametric model to the most informative of the three conditionalized response distributions (see Modeling below).

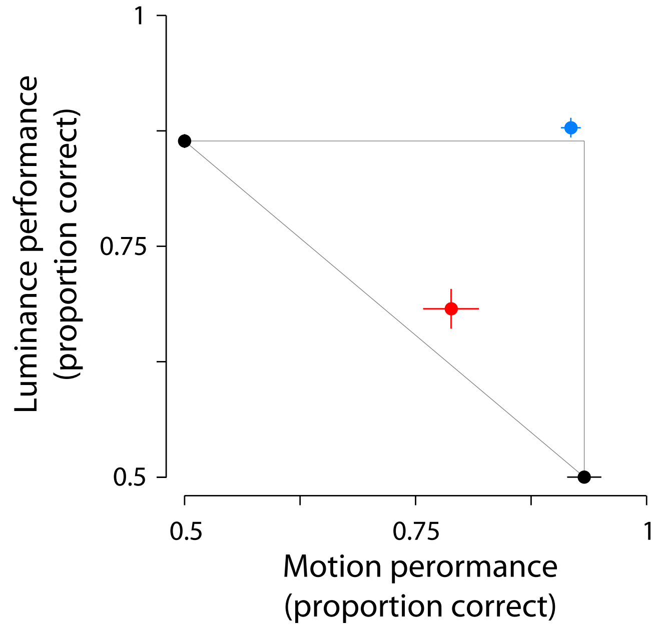

Figure 2. Performance on the motion task (abscissa) is plotted against performance on the luminance task (ordinate). Between-subject average single-task performance is plotted in black on the axes; dual-task performance, when attention was divided within a surface, is plotted in blue; dual-task performance, when attention was divided between-surfaces, is plotted in red. Error bars encompass ±1 standard error of the mean between-subjects (n = 5). Gray lines extend from the single-task performance levels to aid in their comparison to dual-task performance levels.

Error bars

When plotting averages across subjects, error bars encompass ±1 standard error of the mean (n = 5). When plotting individual subject data, we resampled our data (with replacement) 10,000 times, calculating the sample mean after each iteration (Wichmann & Hill, 2001). Error bars enclose ±1 standard deviation of the sampling distribution (±34.14%).

Results

We began by comparing dual-task performance to single-task performance in an attention operator characteristic (AOC) plot (Figure 2) (Sperling & Melchner, 1978). AOC plots are generated by plotting

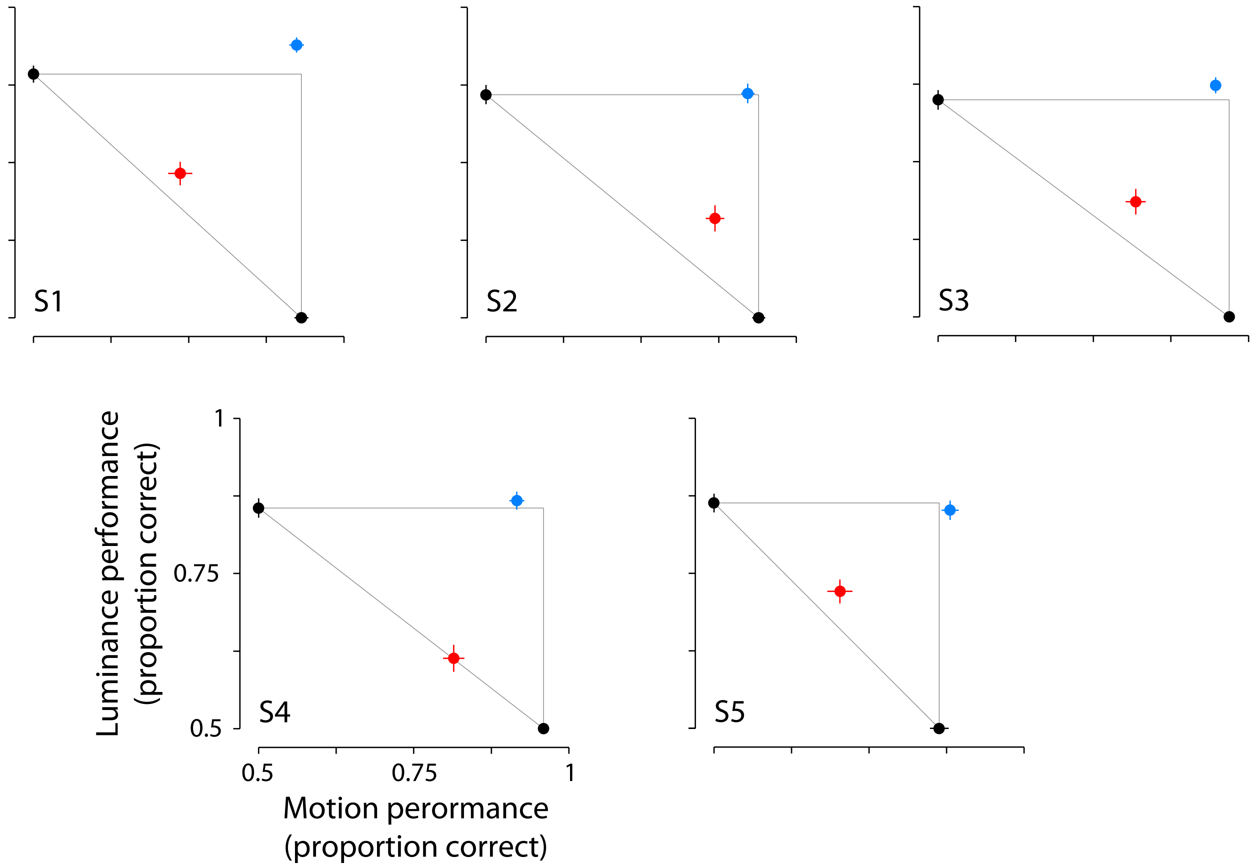

Figure 3. Individual subject performances are summarize by separate AOC plots. Error bars enclose ±1 standard deviation of the bootstrapped sampling distribution (see Methods: Error bars).

dual-task performance levels against one another. To ease comparison between the single-task and dual-task conditions, single-task performance is plotted along the axes. The effect of divided attention within versus between surfaces is readily apparent in the AOC, which shows mean performance across the five subjects.

Detecting a change in both the motion and luminance within a single surface yielded little to no deficit compared to detecting a change in either feature alone. The average within-subject difference between single-task performance and within-surface dual-task performance (0.02 ± 0.01 for the motion task and –0.02 ± 0.02 for the luminance task) was not statistically significant, $t(4) = 1.46$ and 1.45; $p > 0.05$. In contrast, detecting a change in the motion of one surface and the luminance of the other surface resulted in a significant deficit compared to single-task performance levels (–0.14 ± 0.02 and –0.18 ± 0.017). The difference was statistically significant for both tasks, $t(4) = 7.05$ and 10.26; $p < 0.01$. The pattern of dual-task performance was consistent across all five subjects (Figure 3).

To verify that there were no memory or order effects resulting from the order in which the two tasks were performed in the dual-task conditions, we separated performance on the basis of response order (Figure 4). The average within-subject differences between performance in the motion task when motion was queried first versus when color was queried first (–0.013 ± 0.005 for the within-surface condition and –0.009 ± 0.011 for the between-surface condition) was not statistically significant, t(4) = 2.36 and 0.90; $p > 0.05$. Likewise, the differences between performance in the color task when color was queried first versus when motion was queried first (0.00 ± 0.02 and 0.01 ± 0.02) was not statistically significant, $t(4) = 0.12$ and 0.71; $p > 0.05$.

Modeling

We next explored the predictions of two simple nonparametric models at either end of the processing-capacity continuum. The unlimited-capacity sharing model assumes that each task was performed independent of the other task, whereas the all-or-none switching model assumes that only one task can be performed at a given time.

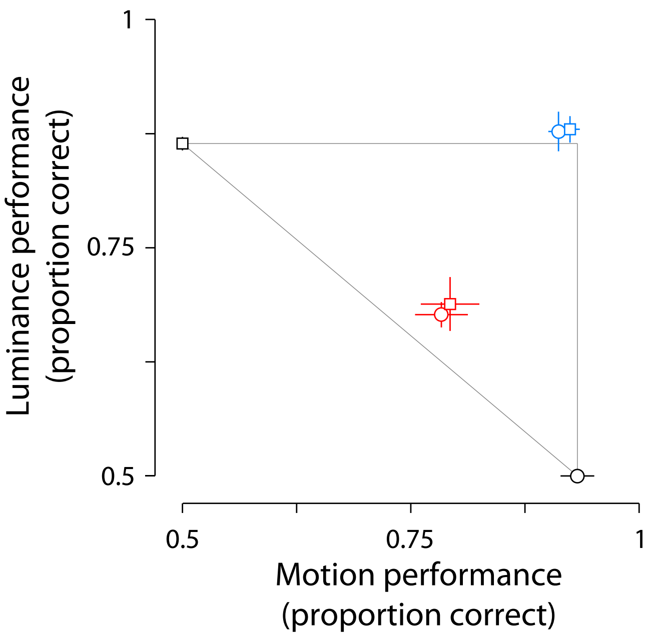

Figure 4. AOC plots, averaged across subjects, after separating dual-task performance on the basis of response order. Squares denote performance over trials in which the luminance task was queried first, and circles denote performance over trials in which the motion task was queried first.

Unlimited-capacity model

There exists a range of models over which the amount of interference between two concurrent tasks varies. At one end of the spectrum resides the unlimited-capacity parallel model, which assumes independent and noninterfering task performance. Given unlimited capacity, the joint probability of two correct responses in the dual-task condition is simply the product of the two single-task performance levels.

$$p(m\,\&\,l)_{dual} = p(m)_{single} \times p(l)_{single} \quad (1)$$

Where m and l represent the motion and luminance tasks, and $p(x)$ represents the probability of a correct response on task x. The unlimited-capacity sharing model predicts that dual-task performance should fall at the independence point at the intersection of the single-task performance levels.

The results from the within-surface dual-task performance fall very close to this intersection for all five subjects (Figure 3), which suggests that there is not a capacity limit for dividing attention across features within a single object.

All-or-none switching model

On the other end of the spectrum resides the all-or-none switching model, which assumes that only one task can be carried out at a given time. Consequentially, the model predicts a negative trial-by-trial correlation in dual-task performance because the observer can only be in one attentional state at a time. When attention is directed to one task, the observer must guess on the other. This model results in two contingency tables, one for each attention state. Since subjects were given no priority instructions, we assumed an even mixture of the two attention states across trials. If the observer attended to motion on half of the trials while guessing on the luminance task, and attended to luminance on the other half of the trials while guessing on the motion task, the joint probability of getting the luminance and motion task correct is:

$$p(m\,\&\,l)_{dual} = \frac{0.5 \times p(m)_{single} + 0.5 \times p(l)_{single}}{2} \quad (2)$$

The all-or-none switching model predicts a trade-off between the two tasks confining dual-task performance along the negative diagonal connecting the two single-task performance levels (see Figures 1 and 2). Equation 2 assumed an equal trade-off between the two tasks resulting in performance halfway between single-task performance and chance. For each of the five subjects (Figure 3), dual-task performance is close to the negative diagonal when attention was divided between surfaces. This seems at first like strong support for the all-or-none switching model for predicting dual-task deficits for dividing attention across surfaces. However, a further analysis shows that this model cannot describe the results.

Rejection of all-or-none switching model: Test for independence

A key feature of the all-or-none switching model is that there should be a negative covariance between trial-by-trial performances, since attention to one task forces the subject to guess on the other. Note that while the AOC plots in Figures 1 and 2 provide a useful graphical summary of performance in divided attention experiments, because performance is collapsed across trials, trial-by-trial covariance cannot be observed in these plots.

The amount of negative covariance predicted by the all-or-none switching model depends on the single-task performance level. Plotted in Figure 5 is the observed dual-task covariance between the motion and luminance task, along with the prediction curve of the all-or-none switching model as a function of single-task performance (Equation 3):

$$cov(PC_{single}) = p(m\,\&\,l) - [p(m\,\&\,l) + p(m\,\&\sim l)]$$

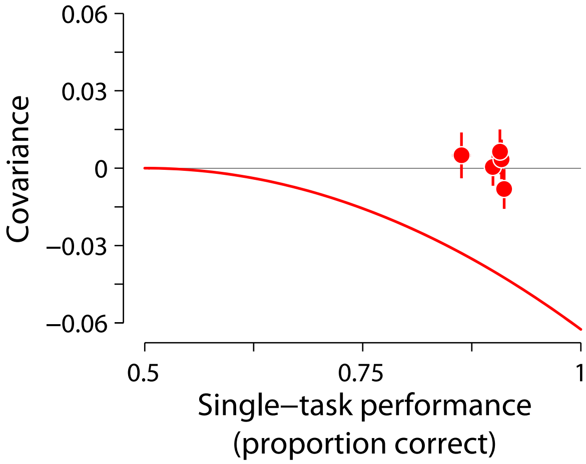

Figure 5. Single-task performance is plotted on the abscissa, and trial-by-trial covariance between the motion and luminance task is plotted on the ordinate. The red data points correspond to observed single-task performance and dual-task covariance for when attention was divided between surfaces. To reduce the dimensionality, single-task performance was averaged across the two tasks. Error bars enclose ±1 standard deviation of the bootstrapped distribution (horizontal error bars fall within the extent of the data points). The red curve traces the covariance predicted by the all-or-none switching model as a function of single-task performance.

$$\times[p(m\,\&\,l) + p(\sim m\,\&\,l)] \tag{3}$$

where $p(m\ \&\ l)$ corresponds to the joint probability of correct response on both task described by Equation 2. To reduce the dimensionality of the space we set $p(m)_{single}$ equal to $p(l)_{single}$ for each point along the abscissa. The "∼" in Equation 3 signifies the probability of an incorrect response—otherwise one minus the probability of a correct response. The between-subject average covariance (0.001 ± 0.003) was statistically indistinguishable from zero, $t(4) = 0.56$; $p > 0.05$. In addition, there was plenty of power to reject the prediction of the all-or-none, switching model. The average within-subject difference between the observer covariance and the model prediction (0.041 ± 0.002) was statistically significant, $t(4) = 16.93$; $p < 0.001$.

Conditionalizing responses on target and distractor transients

Dividing attention across surfaces led to a decrease in performance, but without a corresponding negative trial-by-trial covariance. Instead, some other mechanism besides switching must be causing this performance deficit. One possibility is a phenomenon called *crosstalk* (Navon & Miller, 1987), which is when subjects inadvertently respond to the presence of distractors. For example, subjects may be more likely to respond yes to an upward speed decrement when a speed decrement occurs on the uncued downward moving surface.

To further investigate the crosstalk hypothesis, we conditionalized hits and false alarms on the presence or absence of uncued transients, or *distractors*. If there is crosstalk between channels, then distractors will not be properly filtered, and an increase in hits and false alarms on trials containing distractors would occur. For each task there were three categories of distractors: same feature/different surface, different feature/same surface, and different feature/different surface. For example, when the motion of surface 1 was cued, a distractor transient may occur in the motion of surface 2, the color of surface 1, and/or the color of surface 2.

To begin, we considered distractors within the same feature dimension (Figure 6A, B), e.g., for the motion task, a downward motion transient on trials where upward motion is cued (Figure 6A). The presence of these distractors increased the proportion of yes responses, and therefore both the proportion of hits and false alarms (solid points)—for both tasks and across all conditions—relative to trials containing no distractors (open points). For the motion task (Figure 6A), the average within-subject increase in proportion of yes responses (0.06 ± 0.02 for the single-task, 0.07 ± 0.01 for the dual-task within, and 0.17 ± 0.03 for the dual-task between) was statistically significant across all conditions, $t(4) = 3.40$, 5.06, and 5.93; $p < 0.05$, < 0.01, and < 0.01. For the luminance task (Figure 6B), the average within-subject increase in proportion of yes responses (0.23 ± 0.02, 0.20 ± 0.02, and 0.31 ± 0.03) was statistically significant across all conditions, $t(4) = 13.74$, 8.42, and 10.13; $p < 0.01$.

Based on signal detection theory, (Green & Sweets, 1974), we drew isosensitivity curves (constant d') through the open points to help visualize changes in sensitivity from changes in response bias due to the distractors. For the motion task, the difference in d' (–0.4 ± 0.4, 0.3 ± 0.3, and 0.4 ± 0.3) was not statistically significant, $t(4) = 0.93$, 1.22, and 1.57; $p > 0.05$. For the luminance task, the difference in d' (–0.2 ± 0.3, 0.4 ± 0.3, and 0.5 ± 0.2) was also not statistically significant, $t(4) = 0.61$, 1.54, and 2.10; $p > 0.05$.

Next, we considered distractors within the other feature dimension on the same surface (Figure 6C, D), e.g., for the motion task, a luminance transient in the same surface as the cued motion direction. Luminance transients decreased the accuracy of motion judgments within the same surface across all three cue conditions (Figure 6C). The decrease in sensitivity (by a d' of 1.1 ± 0.2, 1.8 ± 0.2, and 0.6 ± 0.1) was statistically significant across all three conditions, $t(4) = 4.68$, 10.00, and 4.17; $p < 0.01$, 0.01, and 0.05. This suggests that

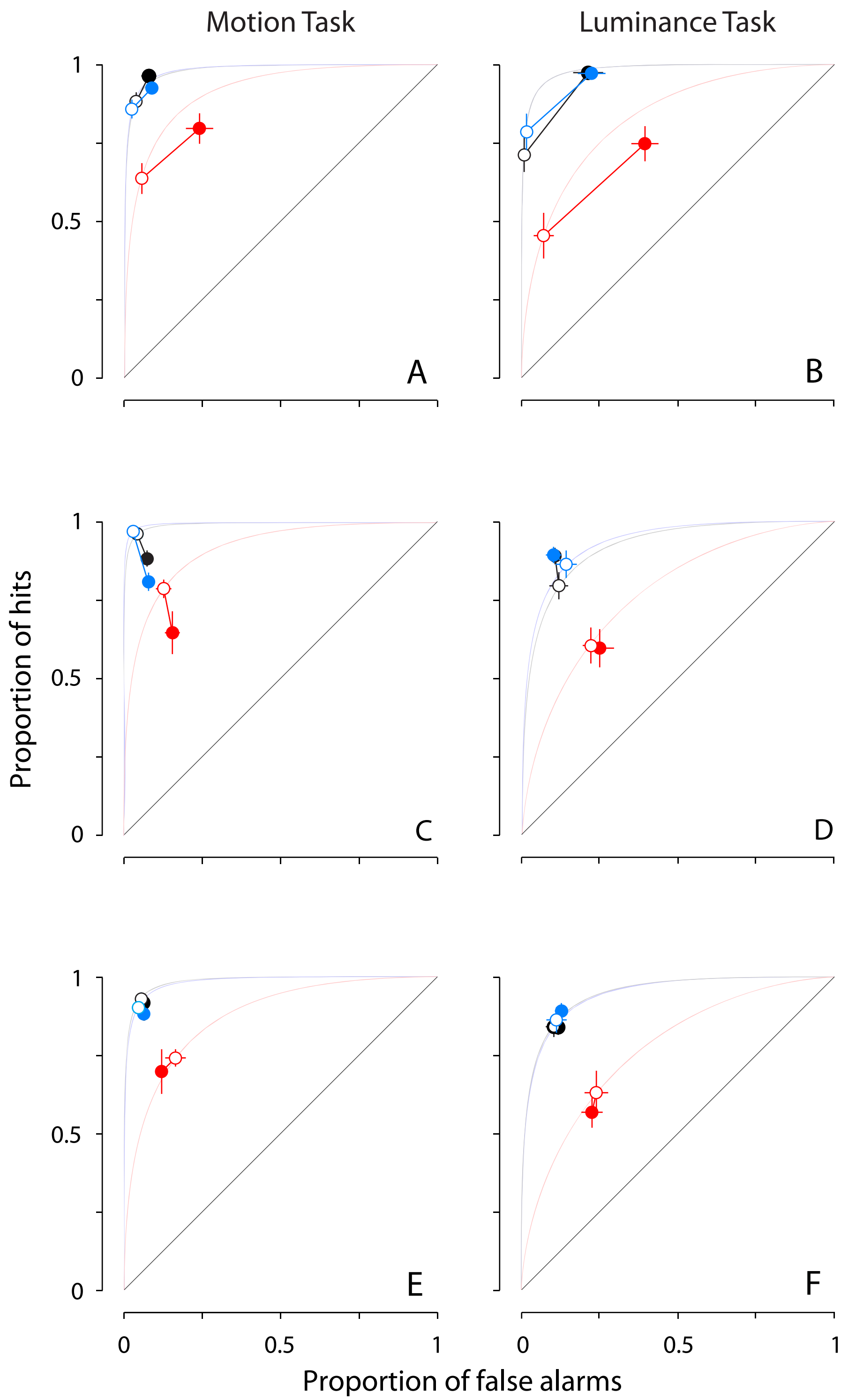

Figure 6. Between-subject average false alarm rate (abscissa) and hit rate (ordinate) contingent on distractor transients. Open points correspond to trials with no distractor transients, and filled points correspond to trials with distractor transients. Isosensitivity curves (assuming zero bias) are drawn through the open points (Green & Sweets, 1974). False alarm and hit rates for (A) the motion task, contingent on motion transients in the other surface; (B) the luminance task, contingent on luminance transients in the other surface; (C) the motion task contingent on luminance transients in the same surface; (D) the luminance task contingent on motion transients in the same surface; (E) the motion task contingent on luminance transients in the other surface; and (F) the luminance task contingent on motion transients in the other surface.

luminance transients masked target-motion transients within the same surface. The complementary effect was not observed in the luminance task (Figure 6D); motion transients did not appear to mask target luminance transient within the same surface. The change in sensitivity to luminance targets when motion distractors occurred in the same surface (–0.5 ± 0.2, –0.3 ± 0.2, and 0.1 ± 0.1) was not statistically significant, $t(4) = 2.68$, 1.54, and 0.83; $p > 0.05$.

Finally, we considered the effect of distractors within the other feature dimension, on the other surface (Figure 6E, F), e.g., for the motion task, a luminance transient in the surface moving in the uncued motion direction (Figure 6E). These distractors had no measurable effect on responses in either task across all three cue conditions. The change in sensitivity for motion transients when luminance distractors occurred (Figure 6E) in the other surface (0.1 ± 0.1, 0.3 ± 0.2, and –0.1 ± 0.2) was not statistically significant, $t(4) =$ 1.21, 0.15, and 0.74; $p > 0.05$. Likewise, the change in sensitivity for luminance transients when motion distractors (Figure 6F) occurred in the other surface (0.1 ± 0.1, 0.00 ± 0.1, and 0.1 ± 0.1) was not statistically significant, $t(4) = 1.12$, 0.03, and 1.46; $p >$ 0.05.

The two main effects captured by the AOC plot (Figure 5) are also present in all of the receiver operator characteristic (ROC) plots (Figure 6). First, the conditionalized responses are similar between the single-task (black) and dual-task within-surface (blue) conditions. Second, when attention was divided between surfaces (red), there was a decrease in sensitivity compared to the other two conditions (i.e., the red points fall closer to the diagonal than the other two conditions). Distractors within the same feature dimension increased the proportion of yes responses across both tasks and in all three conditions (Figure 6A, B). Decrements in luminance masked decrements in motion (decreasing sensitivity) within the same surface (Figure 6C), but not vice versa (Figure 6D). However, luminance events in the other surface had no masking effect on the motion task (Figure 6E).

A limited-capacity sharing model with crosstalk

One way to describe how crosstalk interferes with selection is to imagine that some proportion of the output from the distractor channel is leaked into the output of the target channel. Poor selection can be exemplified by the extent to which the probability of a yes response is greater when: (a) a distractor alone occurred compared to no transients at all, and/or (b) both a target and distractor occurred compared to a target alone.

We formalized this crosstalk concept into a model called *the limited-capacity sharing model with crosstalk*. The term *limited capacity* refers to the fact that we allowed sensitivity to vary freely across conditions, in contrast to the specific limited-capacity model that assumes a fixed rate of information processing (Shaw, 1980). The term *sharing* refers to the assumption that both tasks are performed independently but with limited capacity. The model begins with an encoding stage: each feature is encoded by an independent sensory channel, the output of which is a normally distributed random variable. We assumed that on a transient-absent trial the output of the channel was drawn from a "noise" distribution: a normal distribution with a mean of zero and a standard deviation of one. On transient-present trials the output of the channel was drawn from a "signal" distribution: a normal distribution with a mean greater than or equal to zero and a standard deviation of one. Finally, we assumed that the sensitivity of the two motion/luminance channels were the same (e.g., the sensitivity to upward and downward motion is equivalent).

To illustrate the model, consider the motion-task for a given trial in which upward motion is cued (Figure 7). Figure 7A depicts the probability density functions (PDF) for the outputs of the upward and downward motion channels (above and below respectively). The random variable x_1 denotes the internal evidence for an upward motion transient (target), and the random variable x_2 denotes the internal evidence for a downward motion transient (distractor). In the absence of a transient, the output of either channel is drawn from the noise distribution (with a mean equal to zero). On trials containing a target or distractor transient, the output is drawn from the signal distribution (with a mean shifted from zero). Based on the assumption that attention has no effect on the stimulus encoding stage, the mean of the signal distribution for target and distractor transients are the same. The mean of the signal distribution is equal to the sensitivity (d'). As sensitivity decreases, the overlap between the two distributions increases, making the perceptual discrimination between stimuli present/absent more difficult.

Following stimulus encoding, a decision is based on the output of the cued channel (x_1), plus some amount of leak, or crosstalk, from the uncued channel (x_2). The amount of crosstalk is controlled by a gain term. If selection were perfect then the value of the gain parameter would equal zero and the distractor would have no effect on the decision. If the subject were unable to differentiate the target from the distractor—a complete failure of selective attention—then the gain parameter would equal one. A yes response is made if the pooled output of the two channels is greater than a criterion value (Figure 7B). The proportion of yes responses increases as a function of crosstalk (Figure

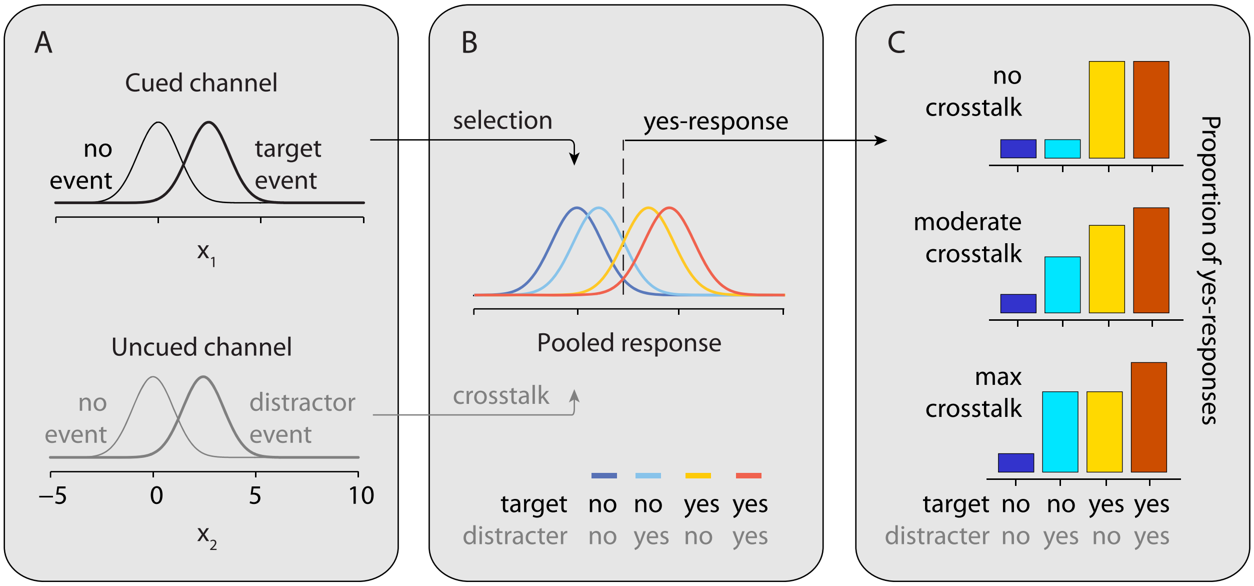

Figure 7. Schematic for the limited-capacity sharing model with crosstalk. (A) The output of two opposing channels (e.g., upward and downward motion-selective channels) is assumed to be normally distributed random variables with unit standard deviation. Trials containing a target transient (above) generate a larger mean response within the cued channel, represented by the shifted PDF. Trials containing a distractor transient (below) generate a larger mean response within the uncued channel, represented by the shifted PDF. A yes/no decision is based on the output of the cued channel plus some leak, or crosstalk, from the uncued channel. There are four possible target/distractor combinations resulting from the two channels. Based on the amount of crosstalk, the means and standard deviations of the four evidence distributions will vary. (B) The four pooled response distributions are shown for a moderate level of crosstalk. The distributions are color-coded as follows: no target or distractor transient (blue), distractor transient alone (cyan), target transient alone (yellow), and both target and distractor transients (red). The dotted gray line represents a possible decision boundary, or criterion, above which a yes response is made. (C) The amount of crosstalk will change the conditional probability of a yes response. Three example distributions are shown for a fixed sensitivity and response criterion given: no crosstalk (top), moderate crosstalk (middle), and max crosstalk (bottom).

7C). If selection were perfect (a crosstalk gain parameter equal to zero), then distractors should have no effect on responses (Figure 7C, top). Given a moderate level of crosstalk (gain = 0.5), distractors will increase the proportion of yes responses (Figure 7C, middle). Given a complete failure of selective attention (gain = 1.0), the proportion of yes responses given a distractor alone will equal the proportion of yes responses given a target alone, and will increase to the combined probability of a yes response when both a target and distractor occur (Figure 7C, bottom).

The model contains three parameters: sensitivity, which defines the mean channel output corresponding to a transient-present trial (signal distribution); a gain term, which controls the amount of leak from the uncued channel; and a decision criterion, which determines how large the pooled output from the two channels must be in order to produce a yes response. We used a maximum likelihood procedure to estimate the parameter values that yielded the greatest probability of generating our observed data set. In order to take full advantage of the information in our data set, we divided trials into four categories based on the pairwise combination of target and distractor transients and tallied the number of yes responses in each category. We then fit the model to these four yes response probabilities. The motion and luminance tasks for each of the three cue conditions were fit separately.

To visualize the model predictions to the data, we replotted in Figure 8 the values from each pair of ROC points from Figure 6A and B on a common axis—the proportion of yes response. Three general patterns are immediately apparent when inspecting the proportion of yes response for each task across the three cue conditions. (a) The distribution of yes responses was nearly identical between the single-task condition and the dual-task, within-surface condition. (b) Crosstalk was more evident for the luminance task than for the motion task across all three cue conditions (compare the proportion of yes responses with and without distractors: cyan vs. blue and red vs. yellow). (c) Performance dropped, and selection errors became more prevalent when attention was divided between-surfaces.

The difference in the observed probability distributions between the single-task and the dual-task, within-surface conditions (Figure 8) was statistically indistinguishable, $\chi^2(3, 4) = 51.81$ and 28.87 for the luminance

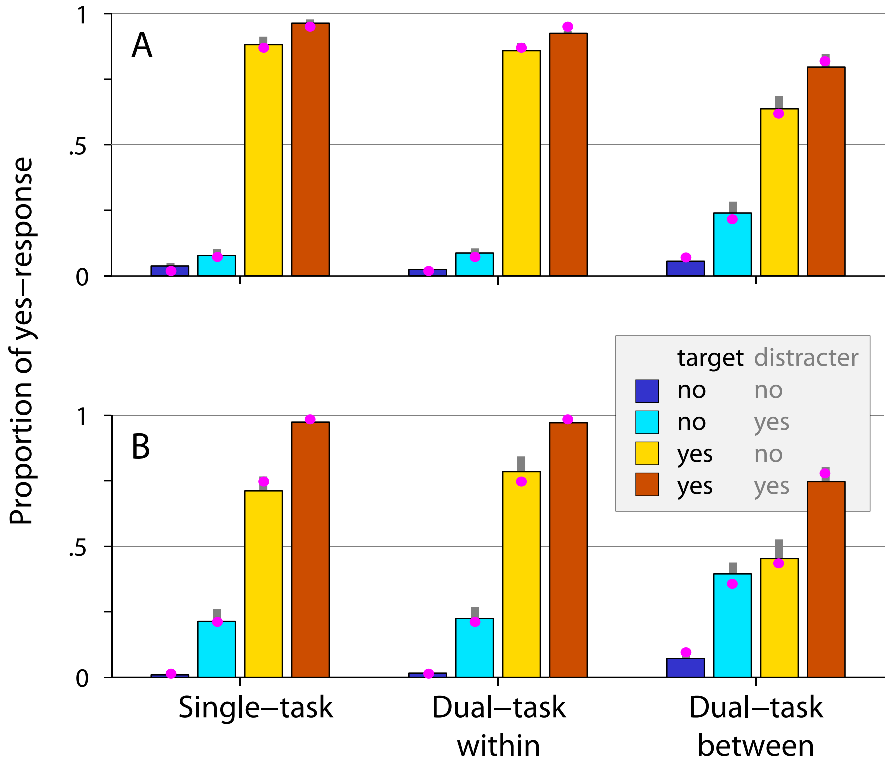

Figure 8. Bars represent the between-subject average proportion of yes responses conditionalized on the four target/distractor combinations (replot of data from Figure 6A, B). Error bars enclose ±1 standard error of the mean. Model predictions are plotted in cyan. Because single-task and dual-task within conditions were simultaneously fit, the cyan points are identical between the two cue conditions. (A) Proportion of yes responses in the motion task given—from blue to orange—no motion transients in either surface; one motion transient in the uncued surface (distractor); one motion transient in the cued surface (target); motion transients in both surfaces (target + distractor). (B) Proportion of yes responses in the luminance task given—from blue to orange—no luminance transients in either surface; one luminance transient in the uncued surface (distractor); one luminance transient in the cued surface (target); luminance transients in both surfaces (target + distractor).

and motion task respectively; $p > 0.05$. Fitting the model separately to these two conditions improved the fit by less than 3% (increase in maximum likelihood) for the motion condition and less than 1% for the color condition. Thus, we reduced our parameters by fitting the combined data for the two conditions (single-task and dual-task within), hitherto referred to as the baseline condition. None of the residual differences between the model predictions and the observed proportion of yes responses were not statistically different from zero ($p > 0.05$ for all 24 t-tests) (Figure 9).

The average parameter values for the baseline and the dual-task, between-surfaces conditions are displayed in Table 2. Considering the baseline condition alone (Table 2, first column), the model describes behavioral performance as follows. First, detection sensitivity was high for both tasks (d' of 3.4 ± 0.2 for the motion task and 3.8 ± 0.2 for luminance task). Second, the crosstalk gain parameter determines how well the subjects were able to select the cued feature and ignore distractor transients within the same feature dimension. The crosstalk gain parameter was significantly greater than zero in both cases, $t(4) = 6.95$ and 17.75; $p < 0.01$, suggesting that even in the baseline condition, subjects were not able to completely filter out distractors. There was more crosstalk in the luminance task than in the motion task (0.50 ± 0.03 vs. 0.17 ± 0.02, for the luminance and motion task, respectively). This difference is reflected in the data by the increase in the false alarm rate when a distractor transient occurred (0.21 ± 0.05 vs. 0.04 ± 0.02, Figure 8 difference between cyan and dark blue bars), and an increase in the hit rate when a distractor transient co-occurred with a target transient (0.26 ± 0.06 vs. 0.08 ± 0.02, Figure 8 difference between orange and yellow bars). Finally, the response criterion determines the trade-off between false alarms and misses. A response criterion equal to half an observer's sensitivity—zero bias—predicts an equivalent false alarm and miss rate. The higher the criterion—a conservative bias greater than zero—the more sensory evidence the observer requires to make a yes response. A conservative observer with a high criterion will commit more misses in order to avoid false alarms. Subjects tended to be conservative in both tasks (bias of 0.45 ± 0.04 and 1.02 ± 0.20, for the motion and luminance tasks, respec-

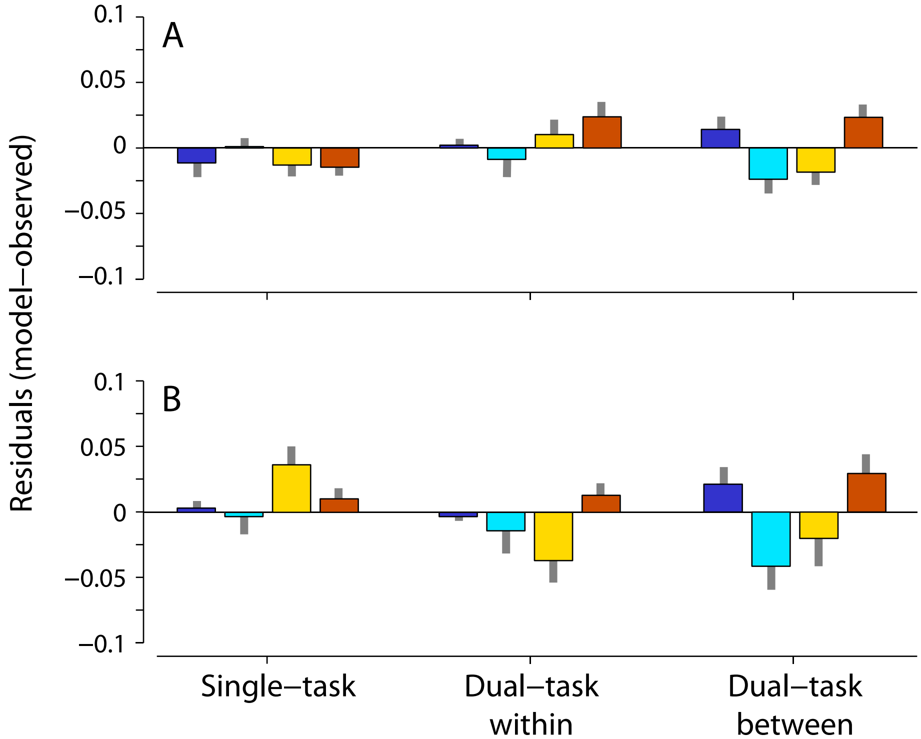

Figure 9. Between-subject residual error; difference scores between the limited-capacity sharing model predictions and the observed proportion of yes responses for the motion task (A) and luminance task (B).

tively), committing fewer false alarms than misses. This parameter is reflected in the data by a very low false alarm rate when no distractor occurred (0.037 ± 0.012 and 0.009 ± 0.004, Figure 8, dark blue bars).

Dividing attention between surfaces is captured in the model by the ratio of the parameter values for the dual-task, between surface and baseline conditions. The $\log_{10}$ of this ratio is tabulated in the third column of Table 2. Dividing attention between surfaces results in: (a) a significant decrease in sensitivity shown by a log sensitivity ratio of −0.21 ± 0.04 for the motion task, and −0.37 ± 0.03 for the luminance task, $t(4) = 4.96$ and 12.97; $p < 0.01$, plotted on the left in Figure 10; (b) a significant increase in crosstalk shown by a log of the crosstalk gain ratio of 0.35 ± 0.09 for the motion task and 0.21 ± 0.04 for the luminance task, $t(4) = 4.09$ and 5.30; $p < 0.05$, plotted on the right in Figure 10. In addition, there was also a conservative shift in bias (corresponding to a change in d' units of 0.26 ± 0.04 and 0.05 ± 0.21) shown by a log criterion ratio of 0.20 ± 0.03 for the motion task and 0.03 ± 0.10 for the luminance task. However, this effect was only significant for the motion task, $t(4) = 5.8$; $p < 0.05$, and not for the luminance task, $t(4) = 0.29$; $p > 0.05$. For the luminance task, 2 of the 5 subjects showed a liberal shift in bias.

An observer may choose to implement a transient detection strategy by ignoring the cue and responding to transients in either surface (e.g., upward or downward speed changes for the motion task; or luminance changes across the red or green dots for the luminance task). When asked to divide attention between surfaces, did the subjects choose to pursue a

	Baseline		Dual-task between		Log parameter ratio	
	Motion	Luminance	Motion	Luminance	Motion	Luminance
Sensitivity	3.4 (0.2)	3.8 (0.2)	2.2 (0.3)	1.6 (0.1)	−0.21* (0.04)	−0.37*(0.03)
Crosstalk gain	0.17 (0.02)	0.50 (0.03)	0.39 (0.07)	0.84 (0.09)	0.35* (0.09)	0.21*(0.04)
Response criterion	2.14 (0.08)	2.92 (0.22)	1.8 (0.1)	1.9 (0.2)	0.20* (0.03)	0.03 (0.10)

Table 2. Average between-subject parameter values for baseline condition (combined single-task and dual-task within), dual-task between surface condition, and the $\log_{10}$ of the ratio between the two (dual-task between divided by baseline). Standard errors of the mean are printed below in parentheses. Parameter values were fit to each subject's data set using a maximum likelihood procedure that maximized the likelihood of the observed conditional proportion of yes responses. The sensitivity and response criterion are in units of d'. The crosstalk gain parameter ranges between zero (no crosstalk) and one (maximal crosstalk). Within-subject t-tests were conducted on the log of the parameter ratios; asterisks denotes a log ratio significantly greater than zero ($p < 0.05$).

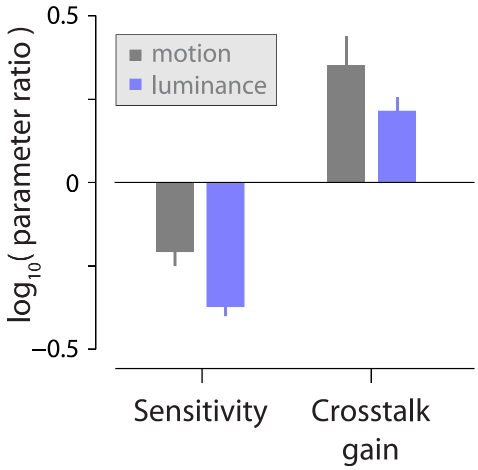

Figure 10. Log_{10} of the ratio between the dual-task between and the combined single/dual-task within parameters (see Table 2) for the motion task (gray) and luminance task (light blue).

transient detection strategy or were they unable to simultaneously select features from competing surfaces? A transient detection strategy would result in poor performance since observers' false alarm rate would equal their hit rate. Such a strategy would place an observer's performance on the negative diagonal in the AOC plots (Figure 3), equivalent to ignoring one of the two cues (as predicted by the all-or-none switching model). Only one subject (S4) demonstrated that level of dual-task deficit. The other four subjects outperformed the theoretical low limit of dual-task performance. In addition, a pure transient strategy would result in a crosstalk gain parameter value of one. The maximum likelihood estimate for the crosstalk gain parameter was below one for both tasks (0.84 ± 0.09 and 0.39 ± 0.07). This suggests that even though selection was poor, subjects were, at the very least, attempting to ignore the distractor transients.

To summarize, the proportion of yes responses conditionalized on target and distractor transients (within the same feature dimension) were statistically identical between the single-task and the dual-task, within surface conditions. Our limited-capacity model with crosstalk adequately fit our data set (Figure 8) with no consistent residual error (Figure 9). Dividing attention between surfaces resulted in a dual-task deficit described by the model as a decrease in detection sensitivity paired with an increase in crosstalk (Figure 10). The increase in crosstalk in the luminance task approached a complete failure of selective attention.

Overall, performance was statistically indistinguishable between the single-task conditions (motion or luminance task) and the dual-task within-surface condition (motion and luminance task). In contrast, dividing attention between surfaces to perform the motion and luminance tasks resulted in a significant dual-task deficit. Performance across the two tasks was statistically independent (uncorrelated), contrary to the predictions of an all-or-none switching model. In addition, distractors within the same feature dimension on the other surface increased the proportion of yes responses, indicative of crosstalk. Although crosstalk was observed in all conditions, it was greatest when attention was divided between surfaces (Figure 6). Although distractors in the other feature on the other surface were successfully filtered (Figure 8), luminance distractors masked motion transients within the same surface (Figure 7). We constructed a limited-capacity sharing model that includes a crosstalk gain parameter to account for crosstalk within the feature dimension. Our model successfully fit the observed proportion of yes responses conditionalized on targets and distractors.

Discussion

We used transparent motion to investigate the capacity limits in divided attention within and between objects. Transparent motion provides a useful stimulus for studying object-based attention because it allows for multiple surfaces to be superimposed, isolating object-based and feature-based selection from spatial selection. Nevertheless, some have suggested that such overlapping displays might be processed by the objects being grouped into different depth planes and selected by 3D spatial attention (see the discussion in Duncan, 1984 and review in Behrmann, Zemel, & Mozer, 1998). This spatial selection hypothesis has two problems. First, the stimuli are 2D with no depth cues. Thus, any 3D interpretation has to come from perceptual organization and not space as specified in the stimulus. Second, studies of attention to 3D space show it is dominated by perceptual organization effects such as attending to a surface rather than local depth cues (He & Nakayama, 1995; Marrara & Moore, 2000). In short, these stimuli have no depth information, and even if they did, it would not help selection. Thus, it is reasonable to assume that object-based and feature-based selection dominates selection for displays with overlapping fields of random dots.

In addition our paradigm possessed two key features that are important for studying attention. First, our stimulus was identical across conditions. Therefore, changes in performance result from capacity-limits in dividing attention rather than sensory encoding effects. Second, the task was held constant between the two divided attention conditions—in both cases the observer performed a motion and luminance task—so changes in performance result from capacity limits in object-based attention rather than task-based effects.

Understanding the effects of divided attention on the processing of multiple features within and between objects is central to models of object-based attention. In this study, we have shown evidence that all features within a relevant surface can be selected with unlimited capacity for the detection of motion and luminance transients. Unlimited capacity for features within an object has also been shown for a wide range of stimuli including: the tilt & texture of a line (Duncan, 1984), the color and shape of a letter (Bonnel & Prinzmetal, 1998), and the orientation, spatial frequency, and color of a Gabor patch (Blaser et al., 2000). This evidence further supports the hypothesis that object-based attention allows unlimited capacity processing of multiple features within an object.

We also showed evidence that dividing attention across two surfaces reduced performance. A deficit in dual-task performance when dividing attention across objects has been reported in a variety of studies across a range of superimposed stimuli (Duncan, 1984; Blaser et al., 2000; Scholl, 2001). Duncan proposed that object-based selection is all-or-none, limiting selection to one object at a time. If selection were all-or-none, then we should expect a negative correlation between a subjects' motion and luminance performance. Instead, we observed no significant negative correlation, and there was enough power (based on high single-task performance levels) to reject the prediction of the all-or-none switching model (Figure 5). Duncan's stimuli involved some degree of spatial segregation which could have contributed to the all-or-none switching effect that he argued for. Consistent with this explanation, subjects' dual-task performance was negatively correlated when they were cued to attend to the shape of one object and the color of a second displaced object (Bonnel & Prinzmetal, 1998). Because we controlled for spatial separation we cannot directly compare our results to those of Bonnel and Prinzmetal.

Valdes-Sosa et al. (2000) also argued that object-based attention was all-or-none using a transparent motion stimulus similar to the one used in this study. However, they manipulated the temporal asynchrony (SOA) between the presentations of the two target probes. It is possible that the first probe exogenously captured attention, leading to the prioritization of information processing from that channel (Shomstein & Yantis, 2004). The time given by the SOA may have encouraged subjects to switch attention between surfaces. We presented targets and distractors simultaneously in order to discourage such a strategy. Our limited-capacity sharing model could be modified to account for differences in task priority, but since we did not manipulate task priority in this study, our model remains ambivalent to the possibility that attention can be flexibly allocated between tasks (but see Bonnel & Prinzmetal, 1998; Sperling & Melchner, 1978, for evidence of flexible allocation).

For each task there were three types of possible distractors. In order to address the influence of these distractors on responses, we analyzed hits and false alarms, contingent on each type of distractor. To begin, we considered transient in different features (luminance distractors during the motion task and vice versa). Distractors within the same surface had an asymmetrical effect on responses. Luminance transients interfered in the detection of motion transients (Figure 6C), but motion transients had no effect on sensitivity to luminance transients (Figure 6D). The masking of motion transients by luminance was observed across all three cue conditions, suggesting that a one-way sensory interaction occurs between luminance and motion when a dot field simultaneously decreases in speed and luminance. Motion detectors (like those found in area MT+) conceived of as spatiotemporal filters (Adelson & Bergen, 1985) may respond to brief luminance changes. A luminance transient produces equal motion energy in all directions and would thus increase the noise across the population of direction-selective neurons, effectively reducing the detectability of motion transients. But contrary to this hypothesis, luminance transients in the other surface had no effect on the sensitivity to motion transients (Figure 8). Regardless of the explanation for this masking phenomenon, it occurred even in the single-task condition and thus does not change our conclusion that processing multiple features within a surface has unlimited capacity.

Next we consider transient in the same feature (i.e., motion distractors on the other surface during the motion task, or luminance distractors on the other surface during the luminance task). Same feature distractors on the other surface had the greatest influence on yes responses (Figures 6 and 10). Whereas the masking effect discussed above may be due to input interference, the failure of selection that occurred within a feature dimension is possibly due to output interference, or crosstalk between channels. Crosstalk and masking are distinguished by the effect of the distractor on the probability of a yes response rather than on the probability of a correct response. Crosstalk was greatest when attention was divided between surfaces as compared to either the single-task or dual-task, within-surface conditions (Figure 8). This failure of selective attention was particularly prevalent in the luminance task—subjects responded with equal probability on trials with a single target or distractor transients. In addition, the probability of a yes response was highest when both a target and distractor transient occurred. This suggests that some portion of the output from the channel encoding information regarding the distractor was leaking into the output of the channel carrying the cued feature information. Crosstalk was present in all conditions (Figure 6), but was most extreme when attention was divided between surfaces.

Could the failure of selective attention that we observed in our data be specific to transparent motion? The answer is probably "no." For example, crosstalk has been reported for simultaneous binaural stimulus presentation (Gilliom & Sorkin, 1974), which, like our paradigm, presents multiple stimuli at the same time. One simple manipulation to explore would be to separate the two surfaces in space, like Vercera and Farah (1994) did with Duncan's (1984) stimuli. Vercera and Farah found no effect of spatial separation on the dual-task deficit between objects, but whether or not we would see an effect on the level of crosstalk using our stimuli remains an open question. A second interesting manipulation that would likely effect the level of crosstalk would be to parametrically vary the heterogeneity between target and distractor features (Lo, Howard, & Holcombe, 2012).

In our experiment, selection was worse for the luminance task than for the motion task. In the baseline conditions, sensitivity was estimated at 3.4 for motion and 3.8 for luminance. In contrast, the crosstalk gain was estimated at 0.17 for motion and 0.50 for luminance. The discrepancy for the two dimensions observed between crosstalk gains seems larger than the discrepancy observed between sensitivities. Thus, this asymmetry seems to be something to take seriously. It is possible that our chromatic feature pairs were less distinguishable than the two motion directions.

Now turn to the larger question of why might selection fail between objects. We suggest two possibilities. First, selecting two objects may result in all of the features of both objects being selected (Chen & Cave, 2006; Egly, Driver, & Rafal, 1994; Yeari & Goldsmith, 2010). In this automatic selection hypothesis, object-based selection is less helpful in selecting the task-relevant information in the between-surface condition compared to the within-surface condition. In particular, none of the irrelevant features within an object are from the same dimension as the relevant feature (e.g., motion-motion). That is not the case for the between-surface condition. A second possibility is that one cannot select two surfaces at once and instead selects the entire stimulus. Again, this allows features from the same dimensions to interfere in the between-surface condition and not in the within-surface condition. These two possibilities might be distinguished by experiments in which there are three surfaces. Can one select two surfaces to prevent interference from a third?

Conclusion

Changes to target features within a cued surface were detected independently and without dual-task cost, consistent with an unlimited-capacity model. By contrast, when the same two target features belonged to different surfaces, detection sensitivity decreased and selection errors increased. Subjects were worse at selecting the cued feature and instead responded to changes in overall intensity, within the feature-dimension, irrespective of surface. Dividing attention across objects interferes with the ability to filter irrelevant features.

Acknowledgments

This work was supported by a National Institutes of Health Grant (EY12925) awarded to Geoffrey M. Boynton.

Commercial relationships: none.
Corresponding author: Zachary Raymond Ernst.
Email: zernst@uw.edu.
Address: University of Washington, Guthrie Hall, Seattle, WA, USA.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the Perception of Motion. *Journal of the Optical Society of America A*, *2*(2), 1861.

Behrmann, M., Zemel, R. S., & Mozer, M. C. (1998). Object-based attention and occlusion: Evidence from normal participants and a computational model. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(4), 1011–1036. doi:10.1037/0096-1523.24.4.1011.

Blaser, E., Pylyshyn, Z. W., & Holcombe, A. O. (2000). Tracking an object through feature space. *Nature*, *408*(6809), 196–199. doi:10.1038/35041567.

Bonnel, A.-M., & Haftser, E. R. (1998). Divided attention between simultaneous auditory and visual signals. *Perception & Psychophysics*, *60*(2), 179–190. doi:10.3758/BF03206027.

Bonnel, A.-M., & Prinzmetal, W. (1998). Dividing attention between the color and the shape of objects. *Perception & Psychophysics*, *60*(1), 113–124. doi:10.3758/BF03211922.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. doi:10.1163/156856897X00357.

Chen, Z., & Cave, K. R. (2006). When does visual attention select all features of a distractor? *Journal of Experimental Psychology: Human Perception and*

Performance, *32*(6), 1452–1464. doi:10.1037/0096-1523.32.6.1452.

Davis, G., Driver, J., Pavani, F., & Shepherd, A. (2000). Reappraising the apparent costs of attending to two separate visual objects. *Vision Research*, *40*(10–12), 1323–1332. doi:10.1016/S0042-6989(99)00189-3.

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London: Series B: Biological Sciences*, *353*(1373), 1245–1255. doi:10.1098/rstb.1998.0280.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222. doi:10.1146/annurev.ne.18.030195.001205.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*(4), 501–517. doi:10.1037/0096-3445.113.4.501.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology. General*, *123*(2), 161–177.

Gilliom, J. D., & Sorkin, R. D. (1974). Sequential vs simultaneous two-channel signal detection: More evidence for a high-level interrupt theory. *The Journal of the Acoustical Society of America*, *56*(1), 157–164. doi:10.1121/1.1903247.

Green, D. M., & Sweets, J. A. (1974). *Signal detection theory and psychophysics*. Huntington, NY: R.E. Krieger.

He, Z. J., & Nakayama, K. (1995). Visual attention to surfaces in three-dimensional space. *Proceedings of the National Academy of Sciences*, *92*(24), 11155–11159.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219.

Kanwisher, N., & Wojciulik, E. (2000). Visual attention: Insights from brain imaging. *Nature Reviews Neuroscience*, *1*(2), 91–100.

Lamy, D., & Egeth, H. (2002). Object-based selection: The role of attentional shifts. *Attention, Perception, & Psychophysics*, *64*(1), 52–66. doi:10.3758/BF03194557.

Lo, S.-Y., Howard, C. J., & Holcombe, A. O. (2012). Feature-based attentional interference revealed in perceptual errors and lags. *Vision Research*, *63*, 20–33. doi:10.1016/j.visres.2012.04.021.

Marrara, M., & Moore, C. (2000). Role of perceptual organization while attending in depth. *Attention, Perception, & Psychophysics*, *62*(4), 786–799. doi:10.3758/BF03206923.

Navon, D., & Miller, J. (1987). Role of outcome conflict in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(3), 435–448.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442. doi:10.1163/156856897X00366.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1–2), 1–46. doi:10.1016/S0010-0277(00)00152-9.

Shaw, M. (1980). Identifying attentional and decision-making components in information processing. *Attention and performance VIII* (pp. 277–392). Hillsdale, NJ: Eribaum.

Shomstein, S., & Yantis, S. (2004). Configural and contextual prioritization in object-based attention. *Psychonomic Bulletin & Review*, *11*(2), 247–253. doi:10.3758/BF03196566.

Sperling, G., & Melchner, M. (1978). The attention operating characteristic: Examples from visual search. *Science*, *202*(4365), 315–318. doi:10.1126/science.694536.

Treisman, A. M. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, *353*(1373), 1295–1306. doi:10.1098/rstb.1998.0284.

Treisman, A. M., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Attention, Perception, & Psychophysics*, *33*(6), 527–532. doi:10.3758/BF03202934.

Valdes-Sosa, M., Cobo, A., & Pinilla, T. (1998). Transparent motion and object-based attention. *Cognition*, *66*(2), B13–B23.

Valdes-Sosa, M., Cobo, A., & Pinilla, T. (2000). Attention to object files defined by transparent motion. *Journal of Experimental Psychology: Human Perception & Performance*, *26*(2), 488–505.

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, *123*(2), 146–160. doi:10.1037/0096-3445.123.2.146.

Watson, S., & Kramer, A. (1999). Object-based visual selective attention and perceptual organization. *Attention, Perception, & Psychophysics, 61*(1), 31–49. doi:10.3758/BF03211947.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, *63*(8), 1314–1329. doi:10.3758/BF03194545.

Yeari, M., & Goldsmith, M. (2010). Is object-based attention mandatory? Strategic control over mode of attention. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(3), 565–579. doi:10.1037/a0016897.