

## Signal Detection Analyses of Repetition Blindness

Nancy G. Kanwisher  
Harvard University

Jin Woo Kim and Thomas D. Wickens  
University of California, Los Angeles

Three experiments used a signal detection model to demonstrate that repetition blindness (N. Kanwisher, 1987) reflects a reduction in sensitivity ( $d'$ ) for the detection of repeated compared with unrepeated visual targets. In Experiment 1, repetition blindness (RB) was found for rapid serial visual presentation (RSVP) letter sequences, whether the visual targets were specified by category membership (vowels) or as 1 of 2 prespecified letters (e.g., *A* or *O*). In Experiment 2, RB was found to a similar degree even when the 1st critical item was displayed for twice as long as the other list items, although overall performance was considerably improved. Experiment 3 found RB for displays containing just 2 simultaneously presented letters. These results support Kanwisher's (1987) account of RB as a genuine perceptual effect, and rule out alternative accounts of RB as the result of response bias, output interference, or guessing biases.

In repetition blindness (RB), people are considerably less accurate at reporting two visually presented items if they are identical or similar than if they are different (Bjork & Murray, 1977; Egeth & Santee, 1981; Kanwisher, 1987; Kim & Kwak, 1990; Mozer, 1989). For example, Kanwisher, Driver, and Machado (1995) found that participants made more errors in reporting two brief simultaneously presented letters when they were the same (two *X*s) than when they were different (an *X* and an *O*). RB occurs not only for letters, but also for colors (Kanwisher, 1991; Kanwisher et al., 1995), words (Kanwisher, 1987; Bavelier & Segui, 1994), and pictures (Bavelier, 1994; Kanwisher & Yin, 1993), presented either simultaneously or in rapid serial visual presentation (RSVP). Kanwisher and Yin asked participants to name the three pictures appearing in an RSVP sequence (and to report an item twice if they saw it twice) and found that the critical pictures were less likely to be correctly reported when they were identical ("bicycle, telephone, bicycle") than when they were different ("bicycle, telephone, house"). RB occurs for a number of different tasks, including full report of RSVP word lists (Bavelier & Potter, 1992; Kanwisher, 1987) or sentences (Bavelier & Potter, 1992; Kanwisher, 1987), partial report of items from spatial arrays (Bjork & Murray, 1977; Egeth & Santee, 1981; Kanwisher, 1991; Kanwisher et al., 1995), report or

detection of single words presented near threshold (Hochhaus & Johnston, 1996; Humphreys, Besner, & Quinlan, 1988; Marohn & Hochhaus, 1988), misspelling detection (MacKay, 1969), and judgments of the number of letters that appear in a simultaneous spatial array (Mozer, 1989).

Kanwisher (1987, 1991) argued that repetition blindness is primarily of interest because it reveals an important functional dissociation between two different visual processes. In *type recognition* the features and identities of visually presented objects are extracted. In *token individuation* a separate episodic representation is set up for each visual object or event. According to the token individuation account of RB, the properties of repeated items are successfully extracted from the second occurrence, but they do not enter awareness because the visual system does not attribute them to a new thing or event. This dissociation between the extraction of visual features and properties, and the binding of those properties to distinct spatiotemporally defined object tokens may be the functional correlate of the anatomical division between the what (ventral) and where (dorsal) pathways in the visual system (Baylis, Driver, & Rafal, 1993).

Whatever its neuroanatomical underpinnings, the token individuation hypothesis takes RB to be a genuine perceptual effect that occurs on-line during the processing of a stimulus. However, several investigators have recently argued that RB is not a bona fide perceptual effect. Armstrong and Mewhort (1993) claimed that both copies of the repeated item are available in memory, and that RB results from a retrieval failure, not an encoding failure. A somewhat similar argument was made by Whittlesea, Dorken, and Podrouzek (1995) and Whittlesea and Podrouzek (1995), although this work has been criticized on methodological and other grounds (Downing & Kanwisher, 1995). Finally, Fagot and Pashler (1995) argued that RB results from a variety of guessing, memory, and output biases involved in full report from RSVP displays, "rather than any

---

Nancy G. Kanwisher, Department of Psychology, Harvard University; Jin Woo Kim and Thomas D. Wickens, Department of Psychology, University of California, Los Angeles.

This research was supported by National Institute of Mental Health Grant MH45245.

We thank Jennifer Keeler and Liana Machado for running participants and Marvin Chun, Paul Downing, Molly Potter, Larry Hochhaus, and Bruce Whittlesea for helpful comments on the article.

Correspondence concerning this article should be addressed to Nancy G. Kanwisher, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, Massachusetts 02138.

fundamental and surprising characteristic of on-line perceptual processing" (p. 290). In the present article we use signal detection theory in an effort to determine whether these alternative accounts of RB are tenable.

In earlier work aimed at this same question, Park and Kanwisher (1994) argued that RB is not due to guessing or response biases or selective forgetting of repeated items that may have been veridically perceived. These claims were supported by several lines of evidence. First, RB is not strongly influenced by total memory load, and it occurs when as few as two items must be reported, so it seems unlikely that RB is due to forgetting of already-encoded items from short-term memory. Second, robust RB effects remain after conservative guessing corrections have been applied to the data (Kanwisher, 1991; Kanwisher et al., 1995). Further, in one experiment in which participants were asked to rate the confidence of their response (Kanwisher et al., 1995, Experiment 4), RB was strongest when only high-confidence responses (which should contain the fewest guesses) were counted as correct. Finally, differences in response bias between repeated and unrepeated items cannot explain why RB is frequently found in some conditions (e.g., with few intervening items) but not others (e.g., with four or more intervening items), even when these conditions are intermixed in the same experiment.

Although we find the aforementioned arguments compelling, signal-detection theory provides an even stronger way to test between accounts of RB in terms of report biases and accounts (such as the token individuation hypothesis) that posit that RB represents an actual reduction in the information the participant has available about the presence of a repeated versus an unrepeated item. If RB can be shown to reflect a reduction in sensitivity ( $d'$ ) for the detection of repeated versus unrepeated items, then the effect cannot be due to guessing or response biases. Accounts of RB in terms of the selective forgetting of repeated items would be ruled out not by the sensitivity effect per se (which in principle could reflect loss of items from memory), but rather by the fact that such reduced sensitivity for repeated items occurred in a detection task in which forgetting is unlikely to be a factor.

Although it has long been clear that it would be useful to measure RB in terms of sensitivity, it has not been obvious how to apply signal detection theory to an RB task. Simply measuring the sensitivity for detection of a repetition is not helpful in itself. Sensitivity is bound to be low but not zero for displays typical of RB experiments. Low sensitivity to detect repeated items would be indicative of RB only if it were lower than the sensitivity to detect unrepeated items. To obtain the proper comparison, Park and Kanwisher (1994) measured detection sensitivity for the second of two targets, as a function of whether this target was the same as the first target (repeated) or different (unrepeated).<sup>1</sup>

Specifically, in their Experiment 7, Park and Kanwisher (1994) showed participants RSVP sequences of letters and told them that each sequence contained at least one vowel. Their task was to determine whether the sequence also contained a second vowel. Participants were specifically trained on the fact that a second vowel might be a second occurrence of the first vowel in that sequence (e.g., two As)

or a new vowel (e.g., an A and an E). After viewing each RSVP letter sequence, the participant first reported whether the sequence contained one or two vowels, and then gave a confidence rating. Trials with one vowel were treated as "noise" events in the signal-detection model, and the two two-vowel conditions, repeated and unrepeated, were treated as distinct "signal" events. If the detectability of the second vowel in the repeated condition is lower than that of the unrepeated condition, then an RB effect has been demonstrated that cannot be due to response or guessing biases. This is what was found: When only one letter intervened between the two target items, sensitivity ( $d'$ ) for detection of the second target was lower when it was the same as the first target (.4) than when it was different (.7). Yet when three or five letters intervened between the two target items, there was no significant difference between sensitivities to repeated and unrepeated items, consistent with past work (Kanwisher, 1987; Park & Kanwisher, 1994) showing that RB declines as the two critical items get farther apart in the list.

The demonstration that RB can be measured in terms of sensitivity seems to argue that RB cannot be due to guessing or response biases (at least in that experiment). It must be noted, however, that Fagot and Pashler (1995, Experiment 3) have recently reported a failure to find RB as measured by  $d'$  in a similar experiment in which participants searched not for a general category (vowels) but rather for two specific letters (e.g., A and B). There are two reasons why this result does not contradict the earlier Park and Kanwisher (1994) result, as it appears at first glance to do. First, in Fagot and Pashler's experiment, all trials contained two targets, and the participant's task was simply to indicate whether they were AA, AB, BA, or BB. This task reverses the usual requirement (in full-report tasks) of token information for repeated trials but only type information for unrepeated trials. If two As are presented, the participant need only note that no B was present to correctly respond AA. However, if both an A and a B were present, the participant must know the order of targets to respond correctly on that trial. Thus, the token individuation hypothesis does not make any clear predictions for this task because participants need not individuate two tokens of one type to respond correctly in the repeated condition, whereas they must encode precise order (i.e., token) information to respond correctly in the unrepeated condition. Second, quite aside from the problem with Fagot and Pashler's task, their signal-detection analysis is uninformative because their model is inconsistent with any RB and because there are internal contradictions in their fitting procedure (see Appendix A for details).

Even though the findings from Fagot and Pashler's (1995) experiment are therefore not inconsistent with Park and Kanwisher's (1994) original claim, other concerns remain. Perhaps the difficulty in detecting the second target in Park and Kanwisher's experiment was not due to an inability to perceive that item, but rather to an inability to categorize it as a vowel and hence a target. Alternatively, if categorizing letters as vowels or nonvowels takes long enough, it is

<sup>1</sup> A similar technique was independently developed by Arnell and Jolicoeur (1993).

possible that participants may actually have forgotten veridically perceived letters before categorization was completed. According to either of these accounts, Park and Kanwisher's  $d'$  measure of RB would constitute not a failure to perceive repeated items, but rather a failure to remember or correctly categorize them.

Hochhaus and Johnston (1996) report a series of experiments designed to get around these difficulties. They used the "single-frame" paradigm (Johnston & Hale, 1984) to measure RB, in which memory load is low because participants make only a single judgment: whether a masked target word is the same as or different from a posttarget probe word. In five experiments, Hochhaus and Johnston found that sensitivity in this task was strikingly and significantly reduced when the target word was preceded by an identical word than when it was preceded by a different word or a string of Xs. Therefore, they argue that RB *can be* a perceptual phenomenon. Presumably the reason for this rather circumscribed conclusion is a concern that these results do not support any strong conclusions for the majority of RB experiments reported in the past, which involve paradigms quite different from the single-frame paradigm.

Thus it remains to be clearly established that the RB that has been reported in the many past studies involving RSVP sequences (e.g., Bavelier & Potter, 1992; Kanwisher, 1987, 1991) and spatial arrays (Kanwisher, 1991; Kanwisher et al., 1995) represents an actual reduction in the sensitivity for detection of repeated versus unrepeated items.<sup>2</sup> That is the goal of the present article. Further, because sensitivity effects can be compatible with memory-based accounts, it is important to demonstrate these sensitivity effects under conditions in which memory accounts are implausible. Thus, variants of the method devised by Park and Kanwisher (1994) are used in which the tasks are simplified and the memory load is reduced.

Reports by the participants in our experiments raised the possibility that detection of second targets may be essentially an all-or-none process. We therefore analyzed our data with both the standard signal-detection model and a high-threshold model. These models differ in the information the participant has available to make the decision. In the standard model, the decision is based on a continuous random variable having a Gaussian (normal) distribution. In the high-threshold model, graded information is not available; the response is based only on the state of a dichotomous variable. These models are not equivalent, because each constrains the data from a rating-scale experiment in a different way. To ensure that our conclusions do not hinge critically on the model selected, both analyses were performed on the data presented here.

### Experiment 1

This experiment was designed to replicate Park and Kanwisher's (1994, Experiment 7) finding of RB with a sensitivity measure ( $d'$ ), as well as to address two concerns about that experiment. We directly compared a replication of our old experiment in which pairs of vowels were detected with

a task in which the identical stimulus lists were presented but the participant was asked to determine whether each list contained one or two targets from a set of two prespecified letters (e.g., A and E). By rearranging the order of trials from the vowel task to create 10 distinct segments in which the same two target letters were searched in each segment (but different letter pairs searched across segments), we could directly compare RB with the  $d'$  measure for the same stimulus sequences as a function of whether participants were looking for vowel targets or specific-letter targets. If our previous RB result was due to a difficulty of categorizing repeated letters as vowels (or remembering the letters long enough to categorize them as vowels), rather than a difficulty in simply perceiving them, then we should not find RB in the specific-letters task.

A final question addressed in this experiment was whether RB declines with experience at the task. Most previous experiments have not been able to definitively answer this question because any change over the course of the experiment could reflect either a real change in RB or simply a change in guessing strategy or criterion. Using the detection statistics as measures should eliminate these ambiguities. By comparing differences in sensitivity to repeated and unrepeated items in the first versus second half of the experiment, we can determine whether participants quickly learn to see visual repetitions. A result reported recently by Chun (1994) used guessing corrections to remove guessing biases and found that RB reversed with practice to become positive repetition priming, but only when target items were differentiated from distractors by a salient feature cue (color). When targets were defined categorically and appeared in the same color as distractors, then there was no reduction of RB with practice.

### Method

*Participants.* Forty participants took part in the experiment. All were members of the University of California, Los Angeles, community; under 30 years of age; had normal or corrected-to-normal vision; and were native English speakers. Participants were either given course credit or paid for their participation in the experiment.

*Materials and design.* Each stimulus consisted of an RSVP sequence of uppercase white letters displayed on a black back-

<sup>2</sup> A much earlier debate about the status of spatial RB began with Bjork and Murray (1977), who showed that participants' ability to report a briefly presented and postcued letter was lower if there was an identical distractor letter in the same display than if the distractor letter was different, an effect replicated by several other investigators including Santee and Egeth (1982), who named it the "repeated-letter inferiority effect." Although Estes (1982) argued that many of the reported effects could be accounted for in terms of bias alone, Santee and Egeth showed that their results could not. Kanwisher (1991) argued that the inconsistency of results in this early work, which sometimes found a repetition benefit and sometimes a repetition cost (Keren & Boer, 1985), resulted from the fact that only a single letter was cued: RB would only be expected when the distractor letter was encoded before the target.

ground. Participants were asked to determine whether each list contained one or two targets. There were five different stimulus conditions: one single-target condition and four two-target conditions. The single-target condition functioned as the "absent" condition in signal-detection terminology, and was used only to calculate the  $d'$ 's for the four two-target "present" conditions. Thus, the distinction between one- versus two-target trials was not a factor in the experiment, but an aspect of the design necessary for the calculation of  $d'$ 's.

Four independent variables crossed in a  $2 \times 2 \times 2 \times 2$  design characterized the two-target trials. The first was task, which was manipulated between participants: either they were told that their targets were vowels (amid consonant distractors) or they were given a different pair of specific target letters (e.g., *A* and *E*) for each segment of trials. Second, in an effort to examine any effects of practice, the data were scored separately for the first half and second half of the experiment. Third, the two stimulus targets were either the same letter (repeated) or they were two different letters (unrepeated). The fourth stimulus factor was lag: either one or four items intervened between the two target items.

There were 120 trials containing a single target and 240 trials containing two targets. The two-target trials were composed of four cells of 60 trials, one cell for each Lag  $\times$  Repetition combination. Within each of the two unrepeated condition cells, one trial was created for each possible combination of first target (C1) identity (*A*, *E*, *I*, *O*, or *U*), second target (C2) identity (one of four vowels not used as C1), and number of consonants appearing before C1 (two, three, or four). Consonants were never repeated within a trial; the consonant sequence for each trial was created by randomly sampling from the set of all consonants except *F*, *Q*, and *V* (because these letters too closely resembled *E*, *O*, and *U*, respectively). Repeated trials were created using the 60 trial combinations described earlier for the unrepeated conditions except that the identity of C2 was changed to be identical to C1. One-target trials were created by including two trials in each of the 60 possible combinations just described, but substituting a consonant for C2. The result of this design was that C1 could appear in serial positions 3, 4, or 5; C2 could appear in serial positions 5, 6, 7, 8, 9, or 10; and sequences contained a total of 11 letters.

The same stimulus materials were used in the two different task conditions; all that differed were the order of trials and the instructions. For the vowel task condition, the order of the 360 trials was randomized and then divided into 10 segments with 36 trials each. The specific-letter task used the identical stimulus materials but separated the 360 trials into 10 segments of 36 trials, one for each of the 10 possible combinations of two-target letters (*AE*, *AI*, *AO*, *AU*, *EI*, *EO*, *EU*, *IO*, *IU*, and *OU*). For example, in the *AE*-target segment, repeated trials contained either two *As* or two *Es*, unrepeated trials contained one *A* and one *E*, and single-target trials contained either an *A* or an *E*. The segment order as well as the trial order within each segment were randomized.

**Apparatus.** This experiment was run on a Macintosh IIci computer with an Apple color high-resolution RGB monitor. The software used for creating and running the experiments was MacProbe Version 1.5.4 written by Steven Hunt. The experiment was carried out in a room lit by a desk lamp with a 60-watt bulb behind the participant pointed at the wall. Although viewing distance was not fixed, most participants viewed the screen from about 50 cm away. From this distance each letter subtended about .60° of visual angle vertically and horizontally.

**Procedure.** Each trial began when the participant pressed the space bar on the computer keyboard. A plus sign (+) appeared in the center of the monitor for 495 ms as a fixation point, followed by sequence of letters presented for 120 ms each, and then a

number (#) sign, which appeared at the end of the sequence for 225 ms as a mask. After 250 ms, a prompt appeared on the screen saying "One or Two?" The participants made their first response by typing a 1 or 2 on the numeric keypad on the right side of the keyboard. Four hundred and ninety-five ms after the participants response, a new prompt appeared on the screen saying "Confidence? (1 → 3)." Participants then typed a 1, 2, or 3 on the numeric keypad on the right side of the keyboard. At the end of each segment of 36 trials, participants were given feedback on the percentage of correct responses they had on the detection task averaged over the last segment and were encouraged to take a short break. The procedure for participants in the letter task condition differed only in that when the participant pressed the space bar to begin each trial, a cue of two target letters appeared for 1,800 ms at the center of the screen, reminding the participant which target letters he or she should look for on that trial.

Participants were told that "When there are two vowels or target letters in a sequence, they may be different (an *O* and a *U*) or the same (two *As*). You don't need to worry about this—your task is simply to decide for each list whether there were one or two vowels or target letters somewhere in that list." For the confidence rating, participants were asked to press 1 if they were sure, 2 if they thought they were probably right but were not positive, and 3 if they were just guessing.

Participants were first run on 36 practice trials, at the end of which they received feedback on their mean percentage of correct responses over the whole practice test. After the practice trials, participants were run on the 360 test trials. The entire experiment lasted about 50 min.

## Results

For each participant, the number of responses in each of the six response categories (one vs. two targets  $\times$  three confidence levels) was tallied for each of the experimental conditions. In our first analysis we calculated  $d'$ 's separately for each block (so that we could look for any effects of practice). In each of the two blocks, the five subconditions for each participant (four two-target conditions plus one one-target condition) were simultaneously fitted by a signal-detection model to provide  $d'$  statistics for each of the four two-target conditions using the one-vowel condition as a reference. This resulted in eight  $d'$  statistics for each participant, which were then analyzed across participants in a 2 (repeated or unrepeated)  $\times$  2 (Lag1 or Lag4)  $\times$  2 (vowels or letter targets)  $\times$  2 (first half or second block) analysis of variance (ANOVA; see Table 1).

Table 1  
*Mean Sensitivities (Gaussian  $d'$ 's) Calculated Separately for Each Participant as a Function of Task, Block, Lag, and Repetition in Experiment 1*

Block no.	Vowel task		Letter task	
	Repeated	Unrepeated	Repeated	Unrepeated
Block 1				
Lag 1	0.50	0.90	0.68	0.98
Lag 4	1.36	1.31	1.26	1.60
Block 2				
Lag 1	0.64	0.87	0.90	1.17
Lag 4	1.16	1.13	1.35	1.74

There was no significant main effect of block,  $F < 1$ , but there was a significant Block  $\times$  Lag interaction,  $F(1, 38) = 10.4, p < .005$ . This interaction reflects the fact that participants improved slightly across blocks in the Lag 1 condition (in both repeated and unrepeated conditions). This may reflect a slight diminution of the attentional blink (Raymond, Shapiro, & Arnell, 1992) with practice (see *Discussion*). However, because the overall effect of practice on lag effects is not our primary interest in this experiment, and because no other variable interacted with block, we reanalyzed our data across blocks (to increase the number of data points contributing to each  $d'$ ). This was done by first collapsing the raw data (i.e., the number of responses in each of the six response categories for each of the nine subconditions for each participant) across blocks and then recalculating the decision statistics. This operation gave eight  $d'$  statistics for each participant, which we then analyzed across participants in a 2 (repeated or unrepeated)  $\times$  2 (Lag1 or Lag4)  $\times$  2 (vowel or letter task) ANOVA. The top half of Table 2 shows the means of these  $d'$  measures.

The ANOVA revealed significant main effects of lower sensitivity in the repeated than unrepeated conditions,  $F(1, 38) = 23.8, p < .001$ ; higher sensitivity for the longer lags,  $F(1, 38) = 154.9, p < .001$ ; and a marginally significant main effect of higher sensitivity in the letter task than the vowel task,  $F(1, 38) = 3.2, p = .08$ . There were also significant interactions of Repetition  $\times$  Lag,  $F(1, 38) = 5.5, p < .05$ , and Task  $\times$  Repetition  $\times$  Lag,  $F(1, 38) = 12.0, p = .001$ , and a marginally significant interaction of Task  $\times$  Repetition,  $F(1, 38) = 3.6, p = .07$ . Following up on the triple interaction, a partial ANOVA of the Lag 1 data revealed significant RB (i.e., a main effect of lower sensitivity in the repeated condition),  $F(1, 38) = 29.5, p < .001$ , and a marginally significant benefit in sensitivity for the letter task compared with the vowel task,  $F(1, 38) = 3.1, p = .09$ . However, importantly, there was no interaction of Repetition  $\times$  Task,  $F < 1$ , indicating that RB at short lags did not vary with task. A partial ANOVA of the Lag 4 data revealed a significant reduction of performance in the repeated condition,  $F(1, 38) = 6.7, p < .02$ , but no significant main effect of task. In this analysis, however, the Task  $\times$  Repetition interaction was highly significant,  $F(1, 28) =$

12.1,  $p = .001$ . The interaction reflects the fact that for the Lag 4 conditions, RB was significant for the letter task,  $F(1, 19) = 12.9, p = .002$ , but not the vowel task,  $F < 1$ .

The data from this experiment were also analyzed using high-threshold alpha. (This analysis uses the rate of two-target responses in one-target trials to obtain corrected hit rates for the various two-target conditions, using the standard guessing-correction formula.) Each participant provided four alpha values, one for each stimulus type. These were analyzed across participants in a 2 (repeated or unrepeated)  $\times$  2 (Lag 1 or Lag 4)  $\times$  2 (Vowel or Letter) ANOVA (see Table 2, bottom half), as in the  $d'$  analysis in the previous paragraph. The pattern of results from this analysis of alpha was very similar to that found for  $d'$ . Specifically, this ANOVA revealed significant main effects of lower sensitivity in the repeated condition,  $F(1, 38) = 23.8, p < .001$ , and higher sensitivity for the longer lag,  $F(1, 38) = 129.2, p < .001$ , but no main effect of task,  $F < 1$ . There were also significant interactions of Repetition  $\times$  Lag,  $F(1, 38) = 4.9, p < .05$ , and Task  $\times$  Repetition  $\times$  Lag,  $F(1, 38) = 9.1, p = .005$ , but no significant interaction between task and repetition,  $F < 1$ . Following up on the triple interaction, a partial ANOVA of the Lag 1 data revealed significant RB,  $F(1, 38) = 24.2, p < .001$ , and no main effect of task,  $F < 1$ . The interaction of Repetition  $\times$  Task did not reach significance,  $F(1, 38) = 2.08, p > .15$ , indicating that RB at short lags does not vary with task. A partial ANOVA of the Lag 4 data revealed a significant reduction of performance in the repeated condition,  $F(1, 38) = 5.7, p < .05$ , but no significant main effect of task. In this analysis, however, the Task  $\times$  Repetition interaction was significant,  $F(1, 28) = 6.5, p < .05$ . The interaction reflects the fact that for the Lag 4 conditions, RB was significant for the letter task,  $F(1, 19) = 11.8, p = .003$ , but not the vowel task,  $F < 1$ .

As can be seen from the previous analysis, the pattern of statistical results was very similar for the standard Gaussian model and the high-threshold model. This was also true in the next two experiments. For a discussion of which model best fits the data, see Appendix B.

*Discussion*

Several conclusions emerge from this data. First, Park and Kanwisher's (1994) finding of RB for short lags using the  $d'$  measure was replicated: the average  $d'$  to detect a target was .95 for unrepeated letters but only .66 for repeated letters. As we will argue in the General Discussion, we suspect that even this sizable effect may be an underestimate of the true RB effect. Second, the finding of RB in this detection task does not hinge on the choice of model, as very similar results were obtained from the standard Gaussian and the high-threshold analyses. Third, RB occurred to a similar extent whether targets were defined categorically (i.e., as vowels) or by enumeration (e.g., As and Es). Thus, the fact that Park and Kanwisher (Experiment 7) found RB for letters in a detection task but Fagot and Pashler (1995) did not cannot be accounted for by the use of categorical targets

Table 2  
Mean  $d'$ s and Alphas Calculated Separately for Each Participant and Condition in Experiment 1

Lag no.	Vowel task		Letter task	
	Repeated	Unrepeated	Repeated	Unrepeated
	Gaussian $d'$			
Lag 1	0.55	0.87***	0.76	1.04**
Lag 4	1.26	1.21	1.27	1.61**
	High-threshold alpha			
Lag 1	0.34	0.51***	0.43	0.52*
Lag 4	0.66	0.66	0.62	0.74**

Note. The significance of repetition blindness for each cell is also shown, based on a  $t$  test between the repeated and unrepeated conditions in that cell.

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

in the former and specific-letter targets in the latter. Fourth, RB did not diminish with practice from the first to the second block (which contained 180 trials each). These findings strengthen Park and Kanwisher's claim that RB is not an artifact of response or guessing biases, but a genuine failure to detect repeated items.

Concerning the fact that RB did not diminish with practice, it should be noted that we did not provide participants with feedback after each trial. One might therefore wonder if RB would be reduced if either more extensive practice at the task was allowed or if trial-by-trial feedback were given. Although we have not run that experiment, we suspect that it would not be very informative. According to Kanwisher's (1987) account of RB, the repeated item does get recognized as a type, although it is not individuated as a new token. Thus, there is information available in the visual system (a higher type activation level, or its equivalent) that participants could in principle learn to exploit as a cue that a repeated item was presented. In other words, participants might learn that a particularly strong sense of the presence of an *A* is a good indicator that two *A*s were present. For this reason we would not be surprised nor would we be particularly enlightened if RB declined with extensive practice and trial-by-trial feedback. What the current experiment shows is that RB is not a transitory problem that participants have in detecting repetitions, which is overcome after experience with only 180 trials.

We have so far focused on the findings from the Lag 1 conditions, in which only one letter intervened between the two target items. This is because earlier work (Kanwisher, 1987; Park & Kanwisher, 1994) has shown that RB is usually gone by the time three to five items intervene between the two occurrences of a repeated item. However, the results of the present experiment have turned up the unexpected result of significant RB in the letter task when four letters (480 ms) intervened between the two occurrences. This long-lag RB effect is found only for the letter task, not the vowel task. It is not clear why RB is sometimes found at longer lags (see also Chun, 1994) and why in particular this subtle task manipulation might change the RB lag effect. We return to this issue in Experiment 2, which also finds RB at long lags.

Although it was not the primary focus of this experiment, an attentional blink (Raymond et al., 1992) is evident in this experiment. The *attentional blink* refers to a temporary reduction in the ability to detect a second visual target in an RSVP sequence within 200–500 ms of the onset of a first visual target. The greatest reduction in second-target detection occurs for an interstimulus interval (ISI) of about 100 ms between the two targets, and performance gradually recovers over the next 3–400 ms. This effect is evident in the lower performance in the unrepeated Lag 1 condition of the present experiment, compared with the unrepeated Lag 4 condition. Assuming this performance difference is indeed due to an attentional blink, then it is noteworthy that when *d*'s were calculated separately for each block, we found a significant overall improvement in performance in the Lag 1 condition in the second block. This may indicate

a reduction of the attentional blink with practice, despite the lack of feedback (see also Martin & Shapiro, 1994).

## Experiment 2

The results of the previous experiment show that RB can be demonstrated in the signal-detection paradigm. However, one concern is that overall performance was quite low even in the unrepeated condition, with *d*'s of around 1 for a lag of one intervening item. It would be more impressive to demonstrate RB under conditions with higher control performance. Second, it would be useful to determine whether RB is in fact affecting the second occurrence of the repeated item in these experiments. The data obtained in Experiment 1 are equally consistent with a model in which the first occurrence is suppressed or overwritten by the second. Although there have been some reports that RB may occasionally act backwards, with a second occurrence suppressing a first occurrence (Bavelier, 1992; Whittlesea et al., 1995), the more common situation in RB is for the second item to be lost (Park & Kanwisher, 1994). In the next experiment the presentation duration of the first occurrence was doubled. This should make detection of the first target very easy, allowing us to more clearly see effects on perception of the second occurrence. If the RB found in Experiment 1 reflects a loss of the second occurrence, not the first, then RB should be at least as strong in this experiment (Park & Kanwisher, 1994, Experiment 2).

## Method

*Participants.* Twenty new participants from the pool previously described in Experiment 1 were run in the experiment.

*Procedure.* The procedure of the experiment was exactly the same as the vowel task version of Experiment 1 with one modification. All letters in each sequence were presented for 120 ms, with the exception of C1, which was presented for 240 ms. Note that increasing the duration of C1 also necessarily changed the C1–C2 stimulus onset asynchrony (SOA) from 240 ms (Lag 1) and 600 ms (Lag 4) in Experiment 1 to 360 ms (Lag 1) and 720 ms (Lag 4) in this experiment.

## Results

For each participant, the number of responses at each of the six levels was tallied for each experimental condition. The data were analyzed as in Experiment 1, to yield both Gaussian *d*' (Table 3, top) and high-threshold alpha (Table 3, bottom).

The four two-vowel conditions plus one one-vowel condition were simultaneously fitted by a signal-detection model to provide *d*' statistics for each of the four two-vowel conditions using the one-vowel condition as a reference. This resulted in four *d*' statistics for each participant, which were then analyzed across participants in a 2 (repeated or unrepeated)  $\times$  2 (Lag 1 or Lag 4) ANOVA (see Table 3).

The ANOVA revealed significant main effects of repetition,  $F(1, 19) = 51.6, p < .001$ , and lag,  $F(1, 19) = 14.9, p = .001$ . There was also a significant Repetition  $\times$  Lag

Table 3  
Mean  $d'$  and Alpha, Calculated Separately for Each Participant and Condition in Experiment 2

Lag no.	Condition	
	Repeated	Unrepeated
	Gaussian $d'$	
Lag 1	1.06	1.93***
Lag 4	1.55	2.05***
	High-threshold alpha	
Lag 1	0.53	0.80***
Lag 4	0.73	0.84**

Note. The significance of repetition blindness for each cell is also shown, based on a  $t$  test between the repeated and unrepeated conditions in that cell.

\*\*  $p = .01$ . \*\*\*  $p < .001$ .

interaction,  $F(1, 19) = 13.0, p = .01$ . Following up on the interaction, the Lag 1 data revealed significant RB:  $d'$  for the repeated condition was significantly less than  $d'$  for the unrepeated condition,  $F(1, 19) = 42.1, p < .001$ . Unexpectedly, analysis of Lag 4 data also revealed significant RB,  $F(1, 19) = 46.6, p < .001$ . Sensitivity was not significantly different for the unrepeated Lag 1 and unrepeated Lag 4 conditions,  $F(1, 19) = 2.4, p > .10$ , indicating that no significant attentional blink occurred in this experiment.

Note that the significant RB effect for Lag 4 in the present experiment was not found in the vowel task data from Experiment 1, yet the present experiment differed from the vowel task conditions of Experiment 1 only in the longer presentation duration of C1 (and the concomitant increase of C1–C2 SOA from 600 ms to 720 ms at Lag 4). To test the significance of this difference in lag effects, we ran a post hoc cross-experiment 2 (repeated or unrepeated)  $\times$  2 (Lag 1 or Lag 4)  $\times$  2 (Experiment 2 or vowel task from Experiment 1) ANOVA. This analysis revealed significant main effects of experiment,  $F(1, 38) = 19.7, p < .001$ ; repetition,  $F(1, 39) = 59.8, p < .001$ ; and lag,  $F(1, 38) = 69.2, p < .001$ . There were also significant interactions of Repetition  $\times$  Lag,  $F(1, 38) = 33.4, p < .001$ ; Experiment  $\times$  Repetition,  $F(1, 38) = 26.6, p < .001$ ; and Experiment  $\times$  Lag,  $F(1, 38) = 4.8, p < .05$ , but no significant Experiment  $\times$  Repetition  $\times$  Lag interaction,  $F < 1$ . The lack of a significant triple interaction suggests that the difference in tasks between the two experiments does not reliably determine whether RB will occur at long lags or not.

The data from this experiment were also analyzed using high-threshold alpha as described in Experiment 1. This resulted in four alpha values for each participant, which were then analyzed across participants in a 2 (repeated or unrepeated)  $\times$  2 (Lag 1 or Lag 4) ANOVA. This ANOVA revealed significant main effects of repetition,  $F(1, 19) = 54.7, p < .0001$ , and lag  $F(1, 19) = 21.1, p < .0001$ . There was also a significant Repetition  $\times$  Lag interaction,  $F(1, 19) = 7.6, p < .05$ . Following up on the interaction, a  $t$  test of the Lag 1 data revealed significant RB,  $t(19) = 6.59, p < .0001$ . A  $t$  test of the Lag 4 data also revealed significant RB,  $t(19) = 3.91, p = .0001$ . In other words, the high-

threshold analysis produced very similar results to the Gaussian analysis.

A post hoc cross-experiment 2 (Experiment 2 or vowel task from Experiment 1)  $\times$  2 (repeated or unrepeated)  $\times$  2 (Lag 1 or Lag 4) ANOVA was also calculated for the alpha values. This analysis revealed significant main effects of experiment,  $F(1, 38) = 7.1, p < .05$ ; repetition,  $F(1, 39) = 56.6, p < .0001$ ; and lag,  $F(1, 38) = 90.1, p < .0001$ . There were also significant interactions of Repetition  $\times$  Lag,  $F(1, 38) = 20.9, p < .0001$ ; Experiment  $\times$  Repetition,  $F(1, 38) = 8.9, p = .0005$ ; and Experiment  $\times$  Lag,  $F(1, 38) = 8.1, p < .01$ , but no significant Experiment  $\times$  Repetition  $\times$  Lag interaction,  $F < 1$ .

### Discussion

The RB observed in this experiment was at least as strong as that observed in Experiment 1, despite the fact that the first target was displayed for 240 ms. This was an easier task, with control unrepeated performance for the Lag 1 condition much higher ( $d' = 1.9$ ) than it was in Experiment 1 ( $d' = .9$ ). Detection of the first target was made almost trivially easy in this task, allowing participants to focus their processing energies on the detection of the second target. Thus, any difficulties that arose in this experiment are likely to reflect on-line perceptual problems, not failures to remember or report the target item. Therefore, the only plausible account of the very large RB effect observed for the Lag 1 condition in this experiment ( $d' = 1.93$  unrepeated;  $d' = 1.06$  repeated) is that participants were actually less likely to perceive second targets when they were repeated than when they were unrepeated.

Second, the fact that RB is at least as strong when C1 is displayed for twice as long as the other list items is consistent with the standard interpretation that RB primarily affects the second critical item, not the first (Kanwisher, 1987; Park & Kanwisher, 1994). Although in some circumstances RB may affect the first occurrence (Whittlesea et al., 1995), such effects are unlikely to have occurred in the present experiment, which produced very strong RB.

Was RB actually stronger when C1 was presented for 240 ms (in Experiment 2) than when it was presented for only 120 ms (in Experiment 1)? Although the post hoc cross-experiment analysis found a significant Experiment  $\times$  Repetition interaction, the larger  $d'$  differences for Experiment 2 (1.93 unrepeated vs. 1.06 repeated for Lag 1) compared with Experiment 1 (.87 unrepeated vs. .55 repeated for Lag 1) are hard to interpret because the baseline unrepeated performance differs so markedly in the two experiments. If RB is measured as a ratio of repeated to unrepeated performance rather than as a difference between unrepeated and repeated performance (Park & Kanwisher, 1994), then the effect is of similar size in the two experiments. In any event the present results are consistent with Park and Kanwisher's argument that RB cannot result from unfinished processing of the first occurrence, because the effect is at least as strong (if not stronger) when the participant is given more time to process C1.

Two further observations from this experiment are worth noting. First, RB was significant even when four letters intervened between the two critical items—in contrast to most past research, in which RB usually disappeared by lags of four intervening items (Kanwisher, 1987; Park & Kanwisher, 1994; but see also Chun, 1994, Experiment 4a). The longer lasting RB found in this experiment, which was also found in the specific-letter version of Experiment 1, may result from a stronger representation of the first occurrence of the repeated item attained from a 240-ms presentation than from a 120-ms presentation.

A second point is that unrepeated performance did not differ for Lags 1 and 4 in this experiment, suggesting that no attentional blink occurs when more time is allotted to the processing of the first target. Note that in increasing the duration of C1 in this experiment we also increased the C1–C2 SOA, and the longer SOAs at both lags in this experiment would be expected to reduce the size of the attentional blink. However, prior work on the attentional blink (Shapiro & Raymond, 1994; Chun & Potter, 1995) finds that performance continues to improve somewhat up to SOAs of 700 ms or more, so the change in SOA is unlikely to explain the complete elimination of the effect. This result is consistent with the fact that the attentional blink disappears if a blank field is presented immediately after the first target (Raymond et al., 1992; Chun & Potter, 1995), a manipulation that also allows more time for the processing of that item. The fact that RB is undiminished when C1 is presented for twice as long but the attentional blink is completely eliminated under these conditions is consistent with Chun's (1994) argument for a double dissociation of the two phenomena.

### Experiment 3

One of the reasons to think that RB is indicative of a very general property of visual information processing is the wide range of stimuli for which it has been found. The aforementioned results complement Hochhaus and Johnston's (1996) findings for the single-frame RB paradigm and show that the RB that occurs when participants view RSVP letter sequences is a real perceptual effect, not a report-level artifact. However, it is worth checking other RB paradigms as well, to make sure they too reflect actual reductions in sensitivity for the detection of repeated items. In the next experiment we test a different kind of stimulus arrangement in which two briefly presented letters are displayed simultaneously, one to the right and one to the left of fixation. Kanwisher et al. (1995) have observed RB for this kind of stimulus when participants are asked to report both letters, or to do a variety of more complex tasks requiring attention to both letters. They have used guessing corrections to argue that the RB effects that result are not due to response bias or guessing strategies. In the next experiment we apply the vowel-detection task to obtain a sensitivity measure for RB for displays of just two simultaneously presented letters.

### Method

*Participants.* Fifteen new participants from the pool previously described in Experiment 1 were run in the experiment.

*Materials and design.* Each stimulus consisted of a small fixation point at the center of the screen and two white letters, one presented on the left and the other presented on the right side of the screen on a black background. This display was presented briefly and was followed by a pattern mask. Participants were asked to determine whether each display contained one or two targets (vowels). There were 180 experimental trials: 60 containing only one vowel with a consonant as the second letter, 60 containing two identical vowels (repeated), and 60 containing two different vowels (unrepeated). There was just one independent variable for the two-vowel stimuli: repeated versus unrepeated.

Within the unrepeated condition, three trials were created for each possible combination of C1 identity (*A, E, I, O, or U*) and C2 identity (one of four vowels not used as C1). Repeated trials were created using the 60 trial combinations described earlier except that the identity of C2 was changed to be identical to C1. One-target trials were created by using the 60 trial combinations just described but substituting a consonant for C2. The consonants were chosen by randomly sampling from the set of all consonants except *F, Q, and V* (because these letters too closely resembled *E, O, and U*, respectively). For the single-target condition, the presentation side of the target was chosen randomly with the constraint that each target letter appeared on each side of the screen equally as often.<sup>3</sup>

As in Experiments 1 and 2, viewing distance was not fixed, but most participants viewed the screen from about 50 cm away. From this distance, each letter subtended about 1.8° of visual angle vertically and horizontally. The letters were presented at a distance of 9.3° from the center of the screen and were centered vertically.

*Procedure.* Each trial began when the participant pressed the space bar on the computer keyboard. A small white fixation point appeared at the center of the screen and remained on throughout the stimulus and mask sequence. After 495 ms, two white letters were presented, one on the far left and another on the far right for a variable duration (see below). This display was followed by two pattern masks (created from line and semicircle elements) that appeared for 495 ms, covering the locations where the two letters had appeared. After a 250-ms blank period, a prompt appeared on the screen saying "One or Two?" The participants made their first response by typing a 1 or a 2 on the numeric keypad on the right side of the keyboard. Four hundred ninety-five milliseconds after the participants' first response, a new prompt appeared on the screen saying "Confidence? (1 → 3)." Participants then typed a 1, 2, or 3 on the numeric keypad on the right side of the keyboard.

A staircase procedure was applied after every four unrepeated trials and was intended to produce around 75% accuracy for these trials. If participants reported the correct number of target letters on all four of these trials, the presentation duration was reduced by one screen refresh (15 ms). If they reported the correct number of target letters on only two or less of the four trials, the duration was

<sup>3</sup> Because of a counterbalancing error, some letters did not occur exactly equal numbers of times on the left and right in the unrepeated condition. For example, *A* appeared on the left in 10 unrepeated trials and on the right in 14 unrepeated trials. However, we do not think this is a problem because there is no reason to believe that the ease of recognizing individual letters should interact with presentation side. In any case, because each letter appeared equally often in the repeated and unrepeated conditions, the present design does not jeopardize our RB measurements.

increased by one screen refresh. Otherwise no change was made. Duration adjustments were based exclusively on performance for unpeated trials, but the resulting duration was applied to all the conditions. The shortest permitted duration was 30 ms, and the longest was 120 ms. The average display duration for the test trials across all participants was 60 ms.

Participants were instructed to fix their eyes on the center fixation point throughout the stimulus sequence. The same task instructions were used as in Experiment 2. That is, participants first decided each stimulus contained one or two vowels and then gave a confidence rating.

Participants were first run on 60 practice trials in which the duration of the letter presentation was adjusted with the staircase procedure from its initial setting of 105 ms. After the practice trials, participants were run on the 180 test trials. There was no break between the practice and the experiment. The entire experiment lasted about 35 min.

### Results and Discussion

For each participant, the number of responses in each of the six response categories was tallied for each of the three experimental conditions. The two two-vowel conditions and the one one-vowel condition were simultaneously fitted by a signal-detection model to provide  $d'$  statistics for each of the two two-vowel conditions using the one one-vowel condition as a reference, as in Experiments 1 and 2. This resulted in two  $d'$  statistics for each participant, which were then analyzed across participants in a  $t$  test between the repeated and the unpeated conditions. The  $d'$  for the repeated condition (1.28) was significantly lower than the  $d'$  for the unpeated condition (1.67),  $t(1, 14) = 4.52, p < .001$ . The same pattern was obtained when the same data were analyzed using the high-threshold model, with alpha for the repeated condition (.48) significantly lower than alpha for the unpeated condition (.66),  $t(1, 14) = 3.45, p < .001$ .

Thus, RB was once again found with the sensitivity measure:  $d'$  for detection of the second vowel was significantly lower if it was the same as the first than if it was different. This finding serves to confirm Kanwisher et al.'s (1995) claim that the RB effects they observed with simultaneously presented letters were not due to guessing or response biases. It also demonstrates that RB can be observed in a simple perceptual detection task containing only two stimulus items.

### General Discussion

This study used a signal-detection model to demonstrate that RB is a true perceptual effect, not an artifact of guessing or response bias. In all three experiments sensitivity for detection of the second target was lower when it was repeated than when it was unpeated. Guessing and response biases cannot account for the differences we observe between repeated and unpeated conditions because these factors cannot differentially affect the detection measures we used. Further, it is very difficult to explain the RB observed in the present experiments in terms of rapid forgetting of repeated items that were originally veridically

perceived (Fagot & Pashler, 1995; Armstrong & Mewhort, 1993). Although one might argue that repeated letters can be perceived but then forgotten before they can be categorized as vowels (in the vowel task of Experiment 1), three aspects of the present data are very difficult to account for in terms of forgetting. First, in the specific-letter task of Experiment 1 it is implausible that a repeated letter would be initially consciously perceived but forgotten before the participant realizes that it is one of the two specific target letters he or she is searching for. For example, a strategy of covert full report followed by numerosity judgment, which might conceivably be used in the vowel task in Experiment 1, is very unlikely to be used when participants need only monitor for two specific target letters. Second, in Experiment 2 the task was made much easier by doubling the presentation duration of the first target, reducing the processing load and allowing participants to focus their energy on the detection of the second target—yet RB was if anything stronger than in Experiment 1. Finally, in Experiment 3 participants simply had to detect a second target in a display containing only two letters, a task with minimal memory load. Yet RB was again robust. Any attempt to explain these effects in terms of selective forgetting of veridically perceived repeated items would be strained at best.

Thus, our main conclusion from the present experiments is that RB cannot be explained in terms of guessing or response biases, and it is not due solely to selective forgetting of repeated items. However, our results are consistent with perceptual accounts of RB, such as the token individuation hypothesis. As described in the introduction, according to this hypothesis RB results from the failure to assign a distinct episodic representation to the repeated item. When this happens the participant does not become aware of the repeated item, and therefore cannot detect it as a target (in the present tasks) or report it (in full or partial report tasks).

We have argued that the sensitivity measure used in this study is immune from guessing and response bias influences, and is unlikely to reflect memory loss. However, it is important to note that there are several other factors (other than RB) that are likely to differentially affect  $d'$  in the repeated and unpeated conditions in the present experiments. For example, after a briefly presented item is consciously recognized and individuated (i.e., after the stage at which RB occurs), it may be easier to categorize it as a target if the same item has just been categorized as a target a moment before (i.e., if it is a repeated C2) than if a different item has just been categorized as a target (i.e., if it is an unpeated C2). This postperceptual effect would produce an increase in sensitivity for the repeated condition over the unpeated condition. Another factor that may differentially benefit  $d'$  (or alpha) in the repeated condition arises simply from the arithmetic of combining probabilities of two events (in this case, detecting two targets): Any inequality in the probability of detecting two of the alternative targets—due to an inherent difference in the recognizability of the two items, or an unequal division of attention between them—will produce an inevitable advantage for the repeated condition whenever the performance measure requires detection of two targets (as in the present

experiments).<sup>4</sup> Thus, there are at least two different reasons why our sensitivity measurements in the repeated condition may be spuriously elevated compared with the unrepeated condition in the present experiments. On the other hand, we can think of no factors (other than those embodied in the RB hypothesis) that will differentially benefit the unrepeated condition. This means that RB is if anything underestimated in the present paradigm. Thus, any failure to find RB with this paradigm would be ambiguous, but any positive RB finding is particularly strong evidence for the perceptual nature of the effect.

The present experiments also have implications about the mechanisms underlying RB. First, the effect does not diminish significantly after 180 trials of practice. This result is consistent with our claim that RB results from a relatively fixed property of the visual system, rather than a default tendency that can be easily changed with practice. Second, Experiments 1 and 2 found significant RB over a greater number of intervening items than has typically been observed in the past (Park & Kanwisher, 1994; Chun, 1994; but see also Chun & Potter, 1992). In Experiment 1 significant RB was found over four intervening items (ISI = 480 ms, SOA = 600), but only for the specific-letter task, not the vowel task. In Experiment 2, in which the first target was displayed for 240 ms, RB was found over four intervening items (ISI = 480 ms, SOA = 720) in the vowel task. Although the explanation for this long-lag RB is not yet clear, one possibility is that situations in which the representation of the first occurrence is particularly strong produce longer lasting susceptibility to RB. Finally, the fact that RB was if anything stronger when the first target was displayed for twice as long, but the attentional blink was reduced to nonsignificance by this same manipulation, is further evidence for a double dissociation of these two effects (see also Chun, 1994).

<sup>4</sup> Suppose the probabilities of detecting the two targets are  $x$  and  $y$ , that they are detected independently, and that trials of the four types ( $x, x; y, y; x, y; y, x$ ) are equally likely. The probabilities of detecting a target in the repeated and unrepeated conditions are  $(x^2 + y^2)/2$  and  $(xy + yx)/2 = xy$ , respectively. The difference between detection in the repeated and unrepeated cases is  $(x^2 - 2xy + y^2)/2 = 1/2(x - y)^2$ , which is greater than 0 whenever  $x$  and  $y$  are unequal.

## References

- Armstrong, I. T., & Mewhort, D. J. K. (1993). *Repetition deficit in RSVP displays: Encoding failure or retrieval failure?* Manuscript submitted for publication.
- Arnell, K., & Jolicoeur, P. (1993, July). *Development of a repetition blindness task that does not require naming.* Poster session presented at the joint meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science and the Experimental Psychology Society, Toronto, Ontario, Canada.
- Ashby, F. G. (Ed.). (1992). *Multidimensional models of perception and cognition.* Hillsdale, NJ: Erlbaum.
- Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154–179.
- Bavelier, D. (1992). *Phonological repetition blindness.* Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge.
- Bavelier, D. (1994). Repetition blindness between visually different items: The case of picture and words. *Cognition*, *51*, 199–236.
- Bavelier, D., & Potter, M. (1992). Visual and phonological codes in repetition blindness. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 134–147.
- Bavelier, D., & Segui, J. (1994). Repetition blindness between words: Nature of the orthographic and phonological representations involved. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1437–1455.
- Baylis, G., Driver, J., & Rafal, R. D. (1993). Visual extinction and stimulus repetition. *Journal of Cognitive Neuroscience*, *5*, 453–466.
- Bjork, E. L., & Murray, J. T. (1977). On the nature of input channels in visual processing. *Psychological Review*, *84*, 472–484.
- Chun, M. (1994). *Processing deficits in RSVP: The attentional blink and repetition blindness.* Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge.
- Chun, M., & Potter, M. (1992, November). *Interference in detecting multiple RSVP targets: Effects of similarity.* Poster session presented at the 33rd annual meeting of the Psychonomic Society, St. Louis, MO.
- Chun, M., & Potter, M. (1995). A two-stage model for multiple target detection in RSVP. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 109–127.
- Dorfman, D. D., & Alf, E. Jr. (1969). Maximum-likelihood estimation of parameters of signal-detection theory and determination of confidence intervals-rating method data. *Journal of Mathematical Psychology*, *6*, 487–498.
- Downing, P., & Kanwisher, N. (1995). Types and tokens unscathed: A reply to Whittlesea, Dorken, and Podrouzek (1995) and Whittlesea and Podrouzek (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1698–1702.
- Egeth, H. E., & Santee, J. L. (1981). Conceptual and perceptual components of interletter inhibition. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 506–517.
- Estes, W. K. (1982). Similarity-related channel interactions in visual processing. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 353–380.
- Fagot, C., & Pashler, H. (1995). Repetition blindness: Perception or memory failure? *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 275–292.
- Goodman, L. A. (1987). New methods for analyzing the intrinsic character of qualitative variables using cross-classified data. *American Journal of Sociology*, *93*, 529–583.
- Greenacre, M. J. (1984). *Theory and application of correspondence analysis.* London: Academic Press.
- Hochhaus, L., & Johnston, J. C. (1996). Perceptual repetition blindness, effects. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 355–366.
- Humphreys, G. W., Besner, D., & Quinlan, P. T. (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General*, *117*, 51–67.
- Johnston, J. C., & Hale, B. L. (1984). The influence of prior context on word identification: Bias and sensitivity effects. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 243–255). Hillsdale, NJ: Erlbaum.
- Kanwisher, N. (1987). Repetition blindness: Type recognition without token individuation. *Cognition*, *27*, 117–143.

- Kanwisher, N. (1991). Repetition blindness and illusory conjunctions. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 404–421.
- Kanwisher, N., Driver, J., & Machado, L. (1995). Spatial repetition blindness is modulated by attention to color and shape. *Cognitive Psychology*, 29, 303–337.
- Kanwisher, N., & Yin, C. (1993). Repetition blindness for pictures: Evidence for rapid computation of abstract pictorial descriptions. *Investigative Ophthalmology and Visual Science*, 34, 1131.
- Keren, G., & Boer, L. C. (1985). Necessary conditions for repeated-letter inferiority: The role of positional uncertainty. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 601–612). Hillsdale, NJ: Erlbaum.
- Kim, J.-O., & Kwak, H.-W. (1990). Stimulus repetition effects and the dimension–feature distinction in alternative targets. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 857–868.
- MacKay, D. G. (1969). The repeated letter effect in misspelling of dysgraphics and normals. *Perception & Psychophysics*, 5, 102–106.
- Marohn, K. M., & Hochhaus, L. (1988). Different-case repetition still leads to perceptual blindness. *Bulletin of the Psychonomic Society*, 26, 29–31.
- Martin, M. J., & Shapiro, K. L. (1994). *Predictable temporal location of the target attenuates the attentional blink effect*. Manuscript in preparation.
- Mozer, M. (1989). Types and tokens in visual letter perception. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 287–303.
- Park, J. and Kanwisher, N. (1994). Repetition blindness: Effects of timing, load, and stimulus configuration. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 500–519.
- Raymond, J., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18, 849–860.
- Santee, J. L., & Egeth, H. E. (1982). Independence versus interference in the perceptual processing of letters. *Perception & Psychophysics*, 31, 101–116.
- Shapiro, K. L., & Raymond, J. E. (1994). Temporal allocation of visual attention: Inhibition or interference? In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory mechanisms in attention, memory, and language* (pp. 151–188). Boston: Academic Press.
- Whittlesea, B., Dorken, M. D., & Podrouzek, K. W. (1995). Repeated events in rapid lists. Part 1: Encoding and representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1670–1688.
- Whittlesea, B., & Podrouzek, K. W. (1995). Repeated events in rapid lists. Part 2: Remembering repetitions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1689–1697.
- Wickens, T. D. (1989). *Multiway contingency tables analysis for the social sciences*. Hillsdale, NJ: Erlbaum.
- Wickens, T. D. (1992). Maximum-likelihood estimation of a multivariate Gaussian rating model with excluded data. *Journal of Mathematical Psychology*, 36, 213–234.
- Wickens, T. D., & Olzak, L. A. (1992). Three views of association in concurrent detection ratings. In G. Ashby (Ed.), *Multidimensional models for perception and cognition* (pp. 229–252). Hillsdale, NJ: Erlbaum.

## Appendix A

### Comment on the Analysis of Fagot and Pashler's (1995) Detection Model

In Fagot and Pashler's (1995) detection model, the four possible target pairs (*AA*, *AB*, *BB*, and *BA*) are represented by equal-variance uncorrelated bivariate normal distributions arranged in a square configuration (their Appendix B and Figure B1). The participants' responses are based on orthogonal criteria drawn parallel to the sides of this square. This representation is completely orthogonal, with equal differences between target means, no correlation in the distributions, and orthogonal response criteria. The orthogonality prevents it from describing effects such as RB that would be expected to alter one or more of these properties. For example, RB for the second appearance of a target might pull their *AA* stimulus nearer to *AB* and *BB* nearer to *BA*, giving the four distributions a trapezoidal configuration. See Ashby and Townsend (1986) and several articles in Ashby (1992) for bivariate models of this type and Wickens and Olzak (1992) for a discussion of the types of effects that may be found in bivariate discrimination data such as these. Although Fagot and Pashler did not provide the individual-participant data necessary to completely

assess their orthogonality assumption, we fitted a bivariate discrimination model similar to the one that they proposed (see Wickens, 1992) to the aggregate Experiment 3 identification results. This analysis suggests that these assumptions are almost certainly violated.

The second difficulty with the Fagot and Pashler (1995) analysis is that it contains internal inconsistencies. Specifically, they assume that the same value of the criterion applies to both targets in a stimulus. They fitted their model in four parts, once for each of the four stimulus types. However, when the analysis is rotated through all four stimuli, the only consistent representation has the same value of the criterion for each stimulus (indeed, it must be zero). To fit them by separate parameters is inconsistent with Fagot and Pashler's analysis. Their parameter estimates (in their Table 3) are neither identical nor zero. Thus, their estimates of the criteria (as well as those we obtained from our fit to their Experiment 3 data) contradict one of the assumptions on which their model is based.

## Appendix B

## Fitting Gaussian and High-Threshold (Alpha) Models to the Data for the Three Experiments

Very similar results were obtained for all three experiments when our data were analyzed using the Gaussian ( $d'$ ) model and the high-threshold (alpha) model, so our conclusions do not depend on the choice of model. Nonetheless, it is of some interest to ask how well each of these models fits the data. A discussion of this question follows.

As mentioned above, two classes of signal-detection model are appropriate for our data. One class contained continuous-information models based on a Gaussian (normal) distribution (the conventional signal-detection model), and the other class contains all-or-none (high-threshold) models. Each of these classes contains several members. In particular, the Gaussian models can either assume equal variance of the underlying distributions, leading to the familiar  $d'$  statistic measuring detectability, or can let the variance be a free parameter, leading to various other measures of detectability, of which the distance  $\Delta m$  between distribution is the most appropriate here. We fitted these models using maximum-likelihood estimates, following a generalization of the algorithm of Dorfman and Alf (1969). (The fitting procedure proposed by Fagot & Pashler, 1995, could not be used here, in part because it is not appropriate for our data, but more importantly because their procedure implicitly assumes that no RB effect is present. See Wickens & Olzak, 1992, for a discussion of the different interpretations of such bivariate detection data.)

*High-Threshold Models*

The all-or-none models all postulate a single discrete detection state, but differ in the rule by which the participant apportions the responses to the rating categories. The most common representation here is the high-threshold model, in which the detection state occurs with probability alpha in the presence of the signal and never in the presence of a noise event. The high-threshold representation is a natural one for RB data—the detection state corresponds to seeing the second presentation. The model here is similar to the “correction for guessing” in test data. It is possible to test the all-or-none representation directly, without assumptions about the response rule. Essentially this model assumes a single binary latent detection state interposed between signal and response. Goodman (1987) showed that this model is equivalent to a correspondence analysis model (e.g., see Greenacre, 1984; Wickens, 1989). Fitting this model leads to tests of the all-or-none character of the data but does not provide identifiable parameter estimates without further assumptions.

For Experiment 1, the general version of the all-or-none model fit the data substantially better than the equal-variance Gaussian

model (36 of 40 participants fit better with the same number of degrees of freedom; these are not hierarchical models, so they cannot be directly tested against each other). This model fit less well than the unequal-variance Gaussian model (7 of 40 participants, significant by a sign test).

*Gaussian Models*

On average, the unequal-variance model fit these data better than the equal-variance model, the difference being significant at the 5% level for 25 of the 40 participants in Experiment 1. However, estimating the additional variance parameters considerably reduced the stability of the detection measures—the standard error of  $\Delta m$  is about half as large as that of  $d'$ . Within this uncertainty, the substantive differences for the two measures were approximately the same.

The main reason for the better fit of the unequal-variance model was, not surprisingly, the greater variance in the two-target conditions relative to the one-target conditions (1.51 vs. 1.00). This result may reflect the type of mixture implied by a two-state representation. An analysis of variance of the two-target scores showed no difference between the materials conditions ( $F < 1$ ), but differences among the various trial types. The variances were somewhat reduced on the Lag 1 repetition trials relative to no repetition (1.27 vs. 1.48) and possibly increased at Lag 4 (1.75 vs. 1.56), although the latter finding is contaminated by some large and unreliable estimates.

Differences in  $\Delta m$  largely paralleled the  $d'$  estimates reported in the *Results* section for Experiment 1, except for an unexpectedly large value in the repetition Lag 4 condition for 1 participant. In turn, this value arose primarily because that participant did not use many of the rating categories. For such reasons, we feel the  $d'$  statistics are the more reliable indicators of performance under this model.

The results of Experiment 2 were also fitted by the various detection models as in Experiment 1. The general character of these fits was the same as in Experiment 1: better fits for the unequal-variance Gaussian model with greater variability of the latent distributions in the two-target conditions, and superior fits for the general all-or-none model than the equal-variance Gaussian model. Again, all measures of detectability measures gave essentially equivalent results.

Received February 3, 1995  
Revision received May 22, 1995  
Accepted August 4, 1995 ■