Feature bindings endure without attention: Evidence from an explicit recall task

DANIEL A. GAJEWSKI
Michigan State University, East Lansing, Michigan

and

JAMES R. BROCKMOLE
University of Edinburgh, Edinburgh, Scotland

Are integrated objects the unit of capacity of visual working memory, or is continued attention needed to maintain bindings between independently stored features? In a delayed recall task, participants reported the color and shape of a probed item from a memory array. During the delay, attention was manipulated with an exogenous cue. Recall was elevated at validly cued positions, indicating that the cue affected item memory. On invalid trials, participants most frequently recalled either both features (perfect object memory) or neither of the two features (no object memory); the frequency with which only one feature was recalled was significantly lower than predicted by feature independence as determined in a single-feature recall task. These data do not support the view that features are remembered independently when attention is withdrawn. Instead, integrated objects are stored in visual working memory without need for continued attention.

Whereas attention is generally thought to play a role in the binding of features into integrated object percepts (see, e.g., Treisman, 1986, 1999; Treisman & Gelade, 1980; Wolfe & Cave, 1999), the role of attention in the maintenance of feature bindings in visual working memory (VWM) is currently being debated. Determining the fragility of the feature bindings is critical because it differentiates among competing views on the representational format of VWM. One view holds that VWM stores integrated object representations (Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001; see also Irwin & Andrews, 1996). Implicit in this view is the idea that an object is represented in memory as a singular structure that either remains coherent when attention is withdrawn or is entirely lost. A second view holds that features are stored independently and attention is needed to maintain proper associations between them (Wheeler & Treisman, 2002). According to this view, removing attention should sever the appropriate bindings, leaving potentially good memory for the features that were present but poor memory for their correct combinations.

Support for the *object-unit hypothesis* was provided in a series of influential studies employing a change detection paradigm (Luck & Vogel, 1997; Vogel et al., 2001). Briefly presented memory displays containing objects comprising two or more simple features (e.g., colored bars in various orientations) were followed by a blank interval and a test display. In separate blocks, participants were to detect changes in one or either of the two feature dimensions. Accuracy depended on the number of objects present in the display, as opposed to the number of features that needed to be remembered. Critically, feature capacity increased even when objects were defined by multiple features from the same feature dimension (e.g., bicolored squares). If equivalent performance for single-feature and multifeature objects were a result of separate stores for each of the feature dimensions, this increase in the number of features from the same dimension should have hindered performance. Thus, these results suggested that integrated objects are the unit of capacity of VWM.

The *independent-stores hypothesis* is based on two criticisms of the Luck and Vogel (1997; Vogel et al., 2001) experiments. First, Wheeler and Treisman (2002) were unable to replicate the finding that feature capacity increases when the objects are comprised of features from the same dimension (see also Olson & Jiang, 2002; Xu, 2002). Second, because the changes in the experiments of Luck and colleagues involved only a single feature of a single item, performance could have been based on memory for the features that occurred in the display without knowledge of which features occurred together. To pursue these issues, Wheeler and Treisman employed a binding condition in which a pair of features were swapped between two items in the initial display. A memory display containing a yellow square and a red triangle, for example, would be tested with a display containing a red square and

a yellow triangle. A decline in performance in the binding condition was observed relative to performance in conditions analogous to those used by Luck and colleagues. As a result, Wheeler and Treisman concluded that features from different dimensions are stored in separate caches in VWM and that these dimensions have independent capacities. Interestingly, this binding-specific decrement was alleviated when memory was tested using a single probe item rather than the whole display. This difference was explained by the suggestions that the preservation of bindings in VWM requires focused attention, and that the onset of the multi-item test display disrupted attention to the feature bindings to a greater degree than did the onset of the single-item test probe.

In the present study, we pitted the object-unit and independent-stores hypotheses against each other by directly investigating the role of attention in the maintenance of feature bindings in VWM. The fundamental question at stake is whether attention, in addition to creating bound percepts of multifeature objects, is also required to preserve those feature bindings in VWM. One potential approach to testing whether or not attention is required to maintain feature bindings in memory would be to manipulate attention during a memory retention interval in a change detection task. In such a situation, the independent-stores hypothesis would predict that a disruption of attention would also disrupt memory for specific feature bindings. The object-unit hypothesis, on the other hand, would predict that variations of attention would have little impact on memory for feature conjunctions. Through this approach, however, support for the object-unit hypothesis would entail the acceptance of a null hypothesis. Therefore, we developed a methodological approach that independently generated specific predictions for memory performance based on the object-unit and independent-stores hypotheses when attention is disrupted. We then contrasted these predicted outcomes against actual performance.

Our approach to the problem had two primary components. First, we employed an exogenous cue to manipulate attention during the memory delay. Nonpredictive cues presented after the offset of a memory display have been shown to elevate memory for a single feature (e.g., color) at the cued location and to reduce memory at the noncued locations (Schmidt, Vogel, Woodman, & Luck, 2002). Here, we extended this paradigm to test memory for the conjunction of two features (color and shape). In particular, we were interested in memory performance at the noncued locations. Because the cue pulls attention away from these locations, memory for the conjunction of features should be especially disrupted if attention is needed to maintain feature bindings. If integrated objects are the unit of capacity, however, memory for the conjunction of features should be no more disrupted than memory for a single feature.

Second, rather than use change detection to make inferences about the contents of memory, we employed an explicit recall task in which observers reported the feature(s) (color and/or shape) of a probed object.[1] The dominant

approach to characterizing the properties of VWM has been the change detection paradigm, yet we think a methodological change is critical for addressing the binding issue. Although change detection situations are capable of determining whether or not an observer's memory matched an actual percept, when memory is imperfect they cannot determine the specific nature of the mismatch. The advantage of the explicit report procedure is that it provides a direct answer to the critical question of how frequently the two features of an object are remembered together. During recall of a colored shape from a memory display, performance in the explicit report task can be broken down in terms of four possible response events: Participants can recall both features, color only, shape only, or neither of the two features. Thus, when memory fails, the exact nature of the failure can be identified, and predictions made by each of the two hypotheses can be tested.

The object-unit view predicts that because features are bound in memory, when attention is removed observers should be able to remember either both features (perfect object memory) or neither of the two (no object memory); outcomes in which a single feature is recalled should occur only by chance when no memory is preserved. The independent-stores view, however, predicts that when attention is pulled away from items in memory each response outcome is an independent event. As a result, all four response events will occur with a meaningful probability determined by the frequency with which a single feature dimension is retained in memory when conjunctions do not have to be maintained. In the following experiment, we describe the precise methodology used to determine predicted outcomes for each response type under the framework of each hypothesis.

METHOD

Participants and Materials

Twelve Michigan State University undergraduate students viewed arrays of colored shapes and reported features associated with a probed item in a subsequent memory test. Each memory display consisted of a circular array of six items. The items were created by randomly combining six colors and six shapes without replacement. The items subtended 1.1º of visual angle and were positioned every 60º around an imaginary circle with a diameter subtending 6.4º of visual angle at a viewing distance of 57 cm.

Procedure

The experiment was divided into three blocks counterbalanced across participants. In two *single-feature* blocks, the participant reported either the color or the shape of the probed item. In the *both-features* block, the participant reported the color and shape of the probed item. Within each block, a 6 (cue position) × 6 (probe position) × 3 (repetition factor) design yielded 108 trials (18 valid and 90 invalid). Six practice trials preceded each block to familiarize the participant with the task. The cue and probe positions of these practice trials were randomly determined. Trial events are depicted in Figure 1. First, to suppress verbal rehearsal, each trial began with the presentation of two digits that the observer vocally repeated throughout the trial. The participant pressed a button to terminate the digits display, which was then replaced with a fixation cross centered on a gray background. A second buttonpress initiated viewing of the stimuli. The memory display was presented for 187 msec
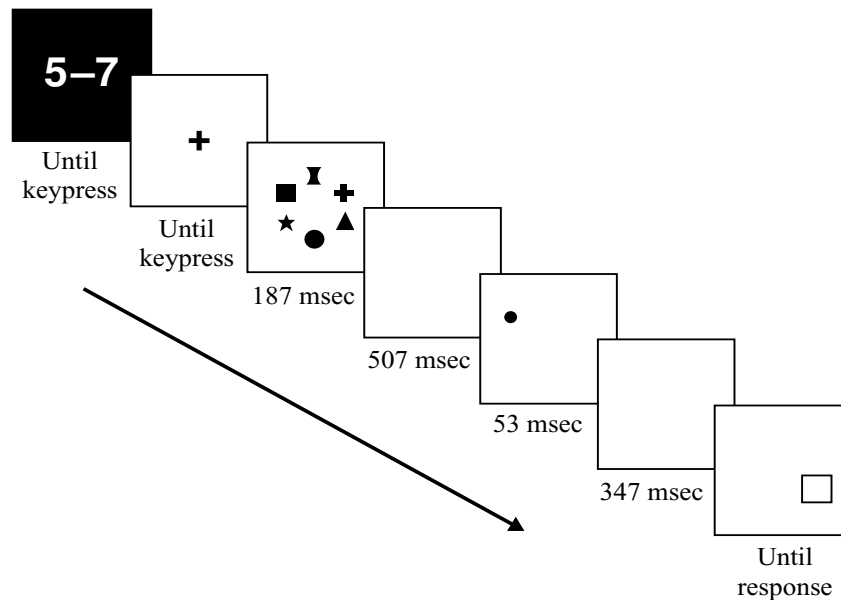
Figure 1. Schematic illustration of the procedure. Each display orthogonally combined the six colors and shapes.

and followed by a 507-msec blank period. A small white dot (the cue) appeared adjacent to one of the six item locations for 53 msec. Following an additional blank period of 347 msec, a gray square frame (the probe) appeared around one of the six locations until the participant pressed the mouse button to indicate his or her readiness to respond. The participant ceased repeating digits and verbally reported the probed item's color and/or shape, depending on the block. The 907-msec interval between the offset of the memory array and the onset of the probe was consistent with memory-to-test intervals typically used in change detection tasks used to investigate feature binding (Luck & Vogel, 1997; Wheeler & Treisman, 2002).

Analysis

In the main analysis of interest, we compared the proportions in which each possible response event was actually observed on the invalid trials of the both-features block with the proportions predicted by the object-unit and independent-stores hypotheses. Predictions for each hypothesis were generated for each individual participant on the basis of his or her performance in each of the single-feature blocks.[2]

Modeling the object-unit hypothesis. For the object-unit hypothesis, memory for individual features was assumed to remain correlated after the removal of attention because of the coherence of the representation. The participants should either remember both features or have to guess both features. The proportion of trials on which the participants reported both features correctly was therefore assumed to be a function of the probability of remembering the item and the probability of guessing both features correctly when the item was not remembered:

$$P = R + [(1 - R) \times G], \quad (1)$$

where P is the probability of reporting both features correctly, R is the probability of remembering the item (i.e., both features), $1 - R$ is the probability of not remembering the item (i.e., remembering neither feature), and G is the joint probability of correctly guessing the color and shape of the item.

Both the proportion of trials on which the participants correctly reported only one feature of the probed object and the proportion of trials on which neither feature was reported correctly were assumed

to be functions of the probability of not remembering the item and of the guessing probabilities associated with each of these events:

$$P = (1 - R) \times G. \quad (2)$$

The predicted probability of recalling one feature correctly is given when G is set equal to the joint probability of correctly guessing one feature and incorrectly guessing the other feature. The predicted proportion of trials on which neither color nor shape is correctly reported is given when G is set equal to the joint probability of incorrectly guessing both features.

To predict values for P in Equations 1 and 2, it was necessary to obtain values for R and G. These values were estimated by performance in the single-feature blocks because, according to the object-unit hypothesis, memory for both features is expected to equal memory for either of the two features alone. Because the probability of correctly guessing the color and shape of the probed object in the both-features block is equal to the product of the probabilities of correctly guessing color and shape individually, the contributions of remembering and of guessing in the single-feature blocks needed to be teased apart.

The first step was to determine the probability of correctly guessing a single feature, G_1. Because each color and each shape was presented exactly once in a display, the probability of correctly guessing a feature of a forgotten item is influenced by the number of features the participant remembers being present at nonprobed locations. Given that there were six to-be-remembered items in the memory display, this can be expressed as

$$G_1 = \frac{1}{6 - 6R_1}, \quad (3)$$

where R_1 is the probability of remembering the probed item in the single-feature blocks.[3] As formulated in Equation 1, the proportion correct in the single-feature blocks, P_1, was assumed to be a function of R_1 and G_1:

$$P_1 = R_1 + [(1 - R_1) \times G_1]. \quad (4)$$

Equation 4 was solved for R_1 by expressing G_1 according to Equation 3 and with P_1 corresponding to the proportion correct across all

cue positions, since guessing is a function of memory for cued and noncued items:

$$R_1 = P_1 - \frac{1}{6}. \tag{5}$$

The single-feature guessing probabilities (G_1s) were computed by returning the resulting R_1 values to Equation 3. To determine the R_1 values for the invalid trials, Equation 4 was then solved for R_1 with the computed single-feature guessing probabilities substituting for G_1 and with P_1 corresponding to the proportion correct on the invalid trials.

Finally, the obtained R_1 and G_1 values were used to generate predictions for the four response events. Because R_1 and G_1 values did not differ between the color and shape blocks ($ps > .40$), means for each participant were used to formulate predictions. Table 1 shows the variants of Equations 1 and 2 that were used to compute the predicted values.

Modeling the independent-stores hypothesis. For the independent-stores hypothesis, memory for color and memory for shape in the both-features block were assumed to be independent events. Performance was also assumed to reflect memory and guessing, but because this hypothesis assumes that features can be remembered and guessed independently, there was no need to tease apart these components of performance. Table 1 shows the equations that were used to compute the predicted values.

RESULTS

The mean proportions of trials (with standard errors) on which each response event occurred, broken down by cue validity and block, are summarized in Table 2. Although there is substantial evidence that onsets capture attention (see, e.g., Jonides & Yantis, 1988; Yantis & Jonides, 1984, 1990), the effect of cue validity was examined in the single-feature blocks to motivate the analyses of interest, which included only the invalid trials. A within-subjects ANOVA was conducted with block feature (color vs. shape) and cue validity (valid vs. invalid) included as factors. Accuracy was higher at the cued location ($M = 0.65$) than at the noncued locations ($M = 0.53$) [$F(1,11) = 7.602$, $MS_e = 0.022$, $p < .05$], and there was no effect of block feature and no interaction (both $ps > .30$).

Object-Unit Model Versus Independent-Stores Model

Because predictions derived from each model were generated for each participant, average model performance was associated with some meaningful variance. Thus, to ensure that the hypotheses made different predictions, an ANOVA was conducted with predictive model (object unit vs. independent stores) and response event (both vs. color only vs. shape only) as within-subjects factors. The proportion of trials on which neither feature was recalled correctly was excluded from all analyses to avoid issues of multicollinearity introduced by the nonindependence of the response events; however, planned comparisons were made for all response events. Figure 2A shows the predicted values for the response events associated with the object-unit and independent-stores hypotheses. Main effects of model [$F(1,11) = 20.79$, $MS_e = .001$, $p < .01$] and response event [$F(2,22) = 10.66$, $MS_e = .025$, $p < .01$] were observed. Critically, the interaction was reliable, indicating that the predicted proportions associated with the response events differed between the models [$F(2,22) = 17.44$, $MS_e = .006$, $p < .001$]. In comparison with the independent-stores hypothesis, the object-unit hypothesis predicts higher proportions of both-correct and neither-correct responses [$F(1,11) = 109.25$, $MS_e = .001$, $p < .001$ and $F(1,11) = 20.84$, $MS_e = .003$, $p < .01$, respectively]. The predicted proportions of color-only and shape-only responses, however, were greater for the independent-stores hypothesis [color only, $F(1,11) = 17.75$, $MS_e = .006$, $p < .01$; shape only, $F(1,11) = 8.97$, $MS_e = .005$, $p < .05$]. To evaluate the predictions of each of the models, these were similarly compared with the data observed on the invalid trials of the both-features block.

Adequacy of the Object-Unit Hypothesis

If objects are the unit of capacity in VWM, the object's features should be remembered either together or not at all (except by chance). Figure 2B shows the observed values and those predicted by the object-unit model. As was expected, the main effect of response event was observed [$F(2,22) = 27.53$, $MS_e = .018$, $p < .001$]. The effect of model was not reliable [$F(1,11) = 1.07$, $MS_e = .008$, $p = .32$]. The model marginally overestimated both-correct response events [$F(1,11) = 3.23$, $MS_e = .011$, $p = .10$], which produced a reliable model × response event interaction [$F(2,22) = 4.26$, $MS_e = .003$, $p < .05$]. Critically, however, the model accurately predicted the proportions of trials on which only one feature was correctly recalled and on which neither of the two features was recalled (all $ps > .32$). The overall pattern of results suggests that items were remembered perfectly or not at all, except as allowed by guessing.

Adequacy of the Independent-Stores Hypothesis

Figure 2C illustrates the observed values and those predicted by the independent-stores hypothesis. Unlike

Table 1
Formulas for Predicting Response Events for the Object-Unit and Independent-Stores Hypotheses

Response Event	Object Unit	Independent Stores
Both	$P = R + \left[(1-R) \times G_1^2\right]$	$P_{\text{BOTH}} = (P_{\text{COLOR}}) \times (P_{\text{SHAPE}})$
Color only	$P = (1-R) \times G_1(1-G_1)$	$P_{\text{COLOR ONLY}} = (P_{\text{COLOR}}) \times (1 - P_{\text{SHAPE}})$
Shape only	$P = (1-R) \times G_1(1-G_1)$	$P_{\text{SHAPE ONLY}} = (1 - P_{\text{COLOR}}) \times (P_{\text{SHAPE}})$
Neither	$P = (1-R) \times (1-G_1)^2$	$P_{\text{NEITHER}} = (1 - P_{\text{COLOR}}) \times (1 - P_{\text{SHAPE}})$

Table 2
Mean Proportions of Trials (With Standard Errors) on Which Each
Response Event Occurred as a Function of Cue Validity and Block

		Response Event							
		Both		Color Only		Shape Only		Neither	
Cue	Block	M	SE	M	SE	M	SE	M	SE
Valid	Color			.68	.04				
	Shape					.62	.05		
	Both	.47	.05	.11	.02	.13	.02	.29	.04
Invalid	Color			.55	.07				
	Shape					.51	.05		
	Both	.34	.05	.13	.02	.13	.01	.40	.07

the object-unit model, the independent-stores model did not fit the data well. An ANOVA revealed main effects of response event [$F(2,22) = 7.52$, $MS_e = .021$, $p < .01$] and model [$F(1,11) = 6.12$, $MS_e = .009$, $p < .05$]. Most importantly, there was a difference between the proportions associated with the observed response events and those predicted by the independent-stores model [$F(2,22) = 4.74$, $MS_e = .008$, $p < .05$]. Although the proportion of trials on which both features were correctly reported did not differ from the proportion predicted by the independent-stores model ($F < 1$), differences among the other three response events were observed. The proportions of observed color-only and shape-only responses were lower than those predicted by the independent-stores model [color only, $F(1,11) = 8.04$, $MS_e = .010$, $p < .05$; shape only, $F(1,11) = 8.88$, $MS_e = .005$, $p < .05$], and the proportion of trials on which neither feature was correctly reported was greater than that predicted by the model [$F(1,11) = 6.14$, $MS_e = .026$, $p < .05$]. This analysis suggests that the features associated with the noncued items were not remembered independently.

DISCUSSION

The goal of the present study was to develop a novel approach to testing competing hypotheses concerning the role of attention in the binding of features in visual working memory. The accomplishment of this goal had two prerequisites: (1) Attention needed to be controlled during the memory interval, and (2) a task was needed that would make different predictions for each of the two views. The first prerequisite was satisfied by the presentation of an exogenous cue within the memory retention interval. The assumption that attention was modulated by the cue was supported by the fact that cue validity influenced the likelihood that the item probed on that trial would be remembered. The second prerequisite was satisfied by the use of an explicit reporting task. On the basis of performance in the single-feature blocks, predictions for the both-features block were generated that differed for (1) the view that bound objects are the unit of capacity and (2) the view that features are remembered independently and attention is needed to maintain feature bindings. The results of the both-features block were unambiguous. The response

frequencies differed from those predicted by feature independence but fit well to the hypothesis that the features of an item are remembered together or not at all, even when attention is called away from the item. This finding is consistent with the view that integrated object representations are the unit of capacity in VWM and inconsistent with the view that attention is needed to maintain feature bindings.

Given the conclusion that the independent-stores hypothesis is untenable, an alternative explanation is required for Wheeler and Treisman's (2002) finding that memory for feature conjunctions was worse when tested using a whole-display probe than when tested with a single-item probe in their binding condition. Although additional experimental work is required to resolve this disparity, one possibility is that their participants adopted different response biases in each condition. When a change was included in a whole-display probe, the majority of the display remained unchanged. Given the high degree of similarity between the memory and test displays in this condition, their participants may have been reluctant to indicate that a change had occurred. On the other hand, when a single-item probe is used to test memory, interference from other items in the display is markedly reduced, potentially providing a less biased measure of memory. Conditions that did not show differences between the whole-display and single-item probes were those in which display changes were created by introducing new features that were not present in the original memory display. As a result, detection rates in the presence of a whole-display test probe may not have differed from those for which a single-item probe was used because the presence of the new features (1) reduced the similarity between the memory display and the whole-display test probe and/or (2) provided a strong cue that a change had occurred.

In conjunction with previous studies, the present research refines our understanding of the relationship between attention and VWM. Attention is required to initially bind object features. If attention is disrupted during this process, binding errors, such as illusory conjunctions (Treisman & Gelade, 1980), can result. In addition, attention influences the transfer of information into VWM. For example, if an exogenous cue precedes the presentation of a to-be-remembered array of single-feature ob-
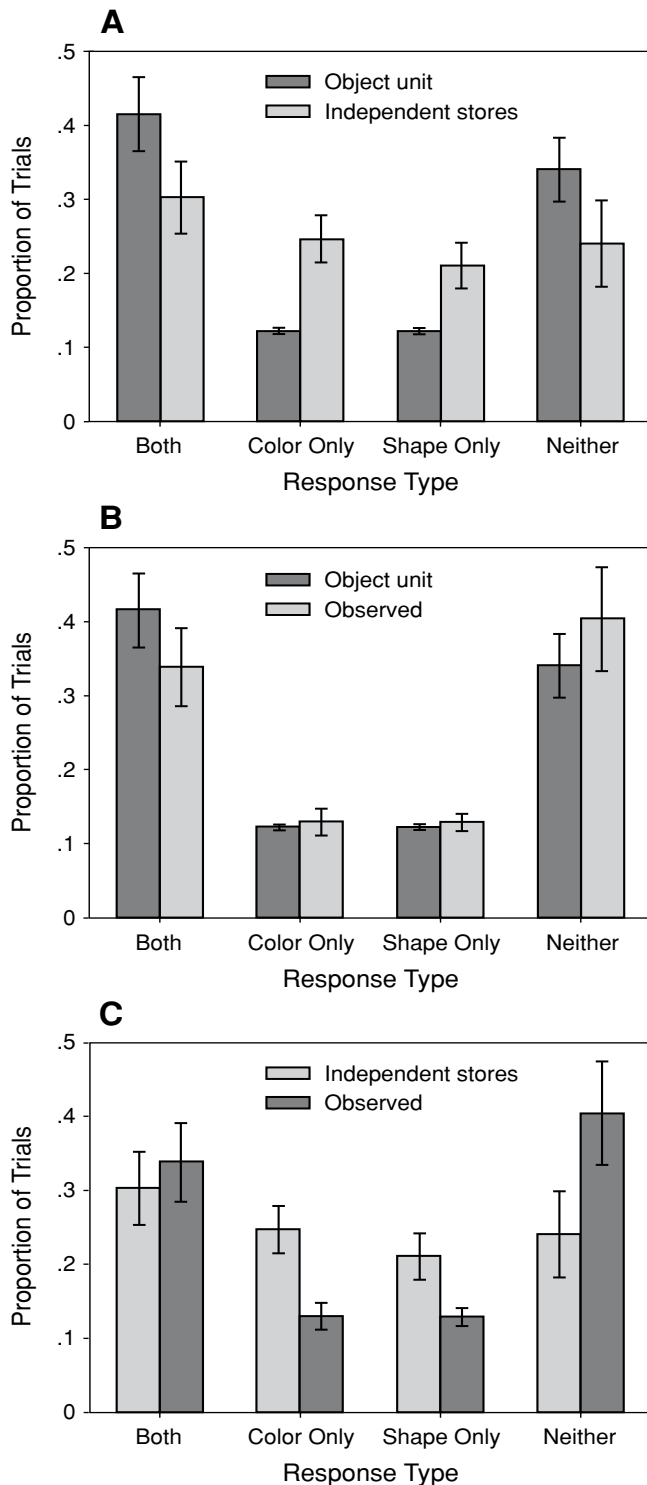
Figure 2. Mean observed and predicted values for the invalid trials of the both-features block. (A) Comparison of predictions generated by the object-unit and independent-stores hypotheses. (B) Comparison of observed performance with performance predicted by the object-unit hypothesis. (C) Comparison of observed performance with performance predicted by the independent-stores hypothesis. All error bars depict standard errors.

jects, the objects near the cue are remembered better than more distal objects (Schmidt et al., 2002). Attention also seems to be required to maintain information in VWM. On the basis of estimations of consolidation speed for similar stimuli, the delayed onset of the cue in the present study provided enough time for the items to be transferred into VWM (Vogel, Woodman, & Luck, in press), but the associated exogenous shift of attention through space during the retention interval nevertheless influenced the probability with which items were retained in memory. Importantly, however, this type of attentional shift does not influence the endurance of feature bindings for those items that are retained. Object representations in VWM appear to be all or none, with entire objects added or deleted from memory.

REFERENCES

IRWIN, D. E., & ANDREWS, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 125-155). Cambridge, MA: MIT Press, Bradford Books.

JONIDES, J., & YANTIS, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, **43**, 346-354.

LUCK, S. J., & VOGEL, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.

OLSON, I. R., & JIANG, Y. (2002). Is visual short-term memory object based? Rejection of the "strong-object" hypothesis. *Perception & Psychophysics*, **64**, 1055-1067.

SCHMIDT, B. K., VOGEL, E. K., WOODMAN, G. F., & LUCK, S. J. (2002). Voluntary and automatic attentional control of visual working memory. *Perception & Psychophysics*, **64**, 754-763.

TREISMAN, A. [M.] (1986). Features and objects in visual processing. *Scientific American*, **225**, 114-125.

TREISMAN, A. [M.] (1999). Feature bindings, attention, and object perception. In G. W. Humphreys, J. Duncan, & A. [M.] Treisman (Eds.), *Attention, space, and action: Studies in cognitive neuroscience* (pp. 91-111). New York: Oxford University Press.

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

VOGEL, E. K., WOODMAN, G. F., & LUCK, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 92-114.

VOGEL, E. K., WOODMAN, G. F., & LUCK, S. J. (in press). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception & Performance*.

WHEELER, M. E., & TREISMAN, A. [M.] (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, **131**, 48-64.

WOLFE, J. M., & CAVE, K. R. (1999). The psychophysical evidence for a binding problem in human vision. *Neuron*, **24**, 11-17.

XU, Y. (2002). Limitations of object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 458-468.

YANTIS, S., & JONIDES, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 601-621.

YANTIS, S., & JONIDES, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 121-134.

NOTES

1. Using a similar approach, Irwin and Andrews (1996) showed that the color and identity of letters are remembered together across eye

movements, but they did not directly investigate the fate of the bindings with the removal of attention.

2. This approach parallels the contrasts used in prior change detection tasks in which performance in multifeature blocks is compared with that in single-feature blocks. In the present study, color and shape were reported no more accurately in the single-feature blocks than they were overall in the both-features block, assuaging concern that observers may employ different memory strategies in each block.

3. Equation 3 assumes imperfect memory for nonprobed positions, since its value is greater than 1 if R_1 exceeds .83 (corresponding to perfect memory for nonprobed locations). However, the maximum observed value for R_1 in the present study was .74.