Psychological Science

Independent Sampling of Features Enables Conscious Perception of Bound Objects

Edward Vul and Anina N. Rich

The online version of this article can be found at:

Published by:

⑤SAGE

On behalf of:

aps
ASSOCIATION FOR
PSYCHOLOGICAL SCIENCE

Association for Psychological Science

Additional services and information for *Psychological Science* can be found at:

Email Alerts: http://pss.sagepub.com/cgi/alerts

Subscriptions: http://pss.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Independent Sampling of Features Enables Conscious Perception of Bound Objects

Edward Vul[1] and Anina N. Rich[2]

[1]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology,
and [2]Macquarie Centre for Cognitive Science, Macquarie University

Abstract

Decades of research suggest that selective attention is critical for binding the features of objects together for conscious perception. A fundamental question, however, remains unresolved: How do people perceive objects, albeit with binding errors (*illusory conjunctions*), when attentional resolution is poor? We used a novel technique to investigate how features are selected to create percepts of bound objects. We measured the correlation of errors (intrusions) in color and identity reports in spatial and temporal selection tasks under conditions of varying spatial or temporal uncertainty. Our findings suggest that attention selects each feature independently by randomly sampling from a probability distribution over space or time. Thus, veridical perception of bound object features arises only when attentional selection is sufficiently precise that the independently sampled features originate from a single object.

People effortlessly perceive scenes comprising numerous objects with many varied features and perceive the correct combination of colors, shapes, and motion that make up these objects (e.g., a red car driving north, a blue bicycle going south). Each object is seen as a cohesive whole, despite the fact that different features (e.g., color, shape, motion) are processed in anatomically segregated parts of the brain (Livingstone & Hubel, 1988).

Although usually observers successfully perceive the correct conjunctions, when selective attention is diverted or impaired, binding of object features can go awry, causing *illusory conjunctions* of incorrect features (e.g., an object with the color of one item and the shape of another; Robertson, 2003; Treisman & Gelade, 1980). Such illusory conjunctions highlight the challenge known as the *binding problem* (Wolfe & Cave, 1999): How does the brain combine information from different specialized areas to provide the subjective experience of cohesive objects? Although psychophysical and physiological evidence suggests that conjunctions are represented in primary visual cortex (Sincich & Horton, 2005) and are formed without attention (Humphrey & Goodale, 1998) or consciousness (Vul & MacLeod, 2006), conscious perception of objects seems to require feature binding by attention (Treisman, 2006).

Most proposals about how attention binds features together for conscious perception suggest that observers infer which features belong to one object by virtue of their location. Feature-integration theory posits that visual attention conjoins features into object files (Treisman & Gelade, 1980; Treisman & Schmidt, 1982), by directing an attentional spotlight to a spatial location and selecting the features therein. Boolean map theory proposes that perception is mediated by a map that defines locations as either selected or not, with features within the same map being bound together (Huang & Pashler, 2007; Huang, Treisman, & Pashler, 2007). However, these accounts leave a fundamental question unanswered: When the attended location is not precise enough to encompass only one object, how are features selected for conscious perception?

Outside the attended region, multiple features seem to be aggregated though a process of statistical summary (Alvarez & Oliva, 2008, 2009; Chong & Treisman, 2005); however, this process produces averages of features rather than illusory conjunctions. Therefore, some researchers have suggested that individual features within an attended region are randomly chosen for perception (Ashby, Prinzmetal, Ivry, & Maddox, 1996; Huang & Pashler, 2007; Treisman & Schmidt, 1982;

Corresponding Author:
Edward Vul, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Ave. 46-4141, Cambridge, MA 02139
E-mail: evul@mit.edu

Vul, Hanus, & Kanwisher, 2009). According to this account, illusory conjunctions arise because different features (e.g., color and form) are chosen independently (Ashby et al., 1996). This account predicts independent intrusions from different features; for example, the report of a color from one item would not predict a report of the form from the same item. Unfortunately, attempts to demonstrate the independence of feature intrusions using accuracy measures (Bundasen, Kyllingsbaek, & Larsen, 2003; Isenberg, Nissen, & Marchak, 1990; Nissen, 1985) have been controversial because irrelevant task factors could introduce, or eliminate, dependence in the accuracy of different feature reports (Monheit & Johnston, 1994; Wolfe & Cave, 1999; e.g., trials on which subjects blinked and missed all the features would induce dependence in accuracy between reports of two features). Such extraneous sources of dependence must be factored out to assess the independence of feature binding.

In this article, we adopt a general statistical framing of the binding process. Following previous accounts, we assume that subjects assess an object's features by estimating the location of the object and evaluating which features were present in that location. Because there will be some uncertainty in location estimates of both the object and the features, this process amounts to probabilistic inference about the co-occurrence of features with the location of interest (Ashby et al., 1996; Vul, Hanus, & Kanwisher, 2009). In a scene with multiple objects, attention mediates this inference by establishing the location of the relevant object in space and time, thus creating a probability distribution that describes the estimated spatiotemporal location of interest. The claim that features are randomly chosen from within a selected region means that features are sampled from this probability distribution. According to this account, decreased precision of attention amounts to worse estimates—and thus increased uncertainty—about the location of the object. In turn, greater uncertainty about the location of the object will result in more errors in the form of feature intrusions.

Given this statistical framing, we designed a new measure to directly test whether conscious perception of conjunctions corresponds to features independently sampled from a probability distribution over space or time. Our goal was to assess whether feature intrusions for different feature dimensions (e.g., color and form) are uncorrelated, as they would be if these dimensions were independently sampled given location uncertainty. In contrast, correlation of feature intrusions would indicate that different features share a source of error, such as internal noise in the location of the attended region. Instead of looking at the accuracy of different feature reports, we evaluated the spatial positions of the reported features and tested whether these positions were correlated between different feature dimensions. When subjects are asked to report both the color and the identity of a letter cued in space, they do not always report the correct color and letter; rather, they frequently report spatially proximal colors and letters. Our question was, if a subject reports the color to the right of the target,

does this predict that the subject will also report the letter to the right of the target? We answered this question by looking at the correlation in spatial position errors between the two features. This measure enabled us to detect systematic relationships between feature intrusions and therefore to assess the dependence of one feature report on the other while factoring out shared task factors.

To demonstrate that independence, as measured by a lack of correlation, was not due to limitations of memory (Tsal, 1989; Wolfe & Cave, 1999), task demands, or statistical power, we included conditions in which we introduced external noise to the cue, to simulate the possible effects of internal noise. This manipulation made the cue less accurate, so that it effectively pointed to items on either side of the target on some trials. We expected this to cause a systematic relationship in the errors of color and letter identity, because errors in cue position would contribute to the position error of both feature reports. This manipulation allowed us to verify that our method is able to detect correlations of position intrusions across features when we know that they should be present.

Our results show that in both space (Experiment 1) and time (Experiment 3), illusory conjunctions arise from a process that samples features independently: There is no correlation between intrusions in color and intrusions in letter identity. Furthermore, this lack of correlation cannot be ascribed to limitations of memory, task demands, or statistical power, because in both space (Experiment 2) and time (Experiment 4), the external-noise manipulation produced reliably correlated feature intrusions.

Experiments 1 and 2: Binding in Space
Method

Participants. In Experiment 1 (spatial uncertainty), 10 participants (6 female, 4 male; ages 18–40 years) from the Massachusetts Institute of Technology subject pool were paid for participation. In Experiment 2 (spatial noise), 12 participants (10 female, 2 male; ages 18–40 years) from the Macquarie University subject pool were given course credit or were paid for participation.

Materials and design. On each trial, subjects viewed a brief presentation of 26 colored capital letters arranged in a circle (6° diameter) centered at the fixation point and reported the color and letter identity of one item cued as the target (see Fig. 1a). The colors and identities of the five items around the target were unique, which allowed us to identify the spatial or temporal position, relative to the cued item, corresponding to each reported feature. The 26 English letters were presented in a random order in Courier font. Each letter was randomly assigned one of five colors with the constraint that the target letter was the center of a set of five uniquely colored letters. At the viewing distance of approximately 57 cm, the letters subtended approximately 1.3° of visual angle at about 6°
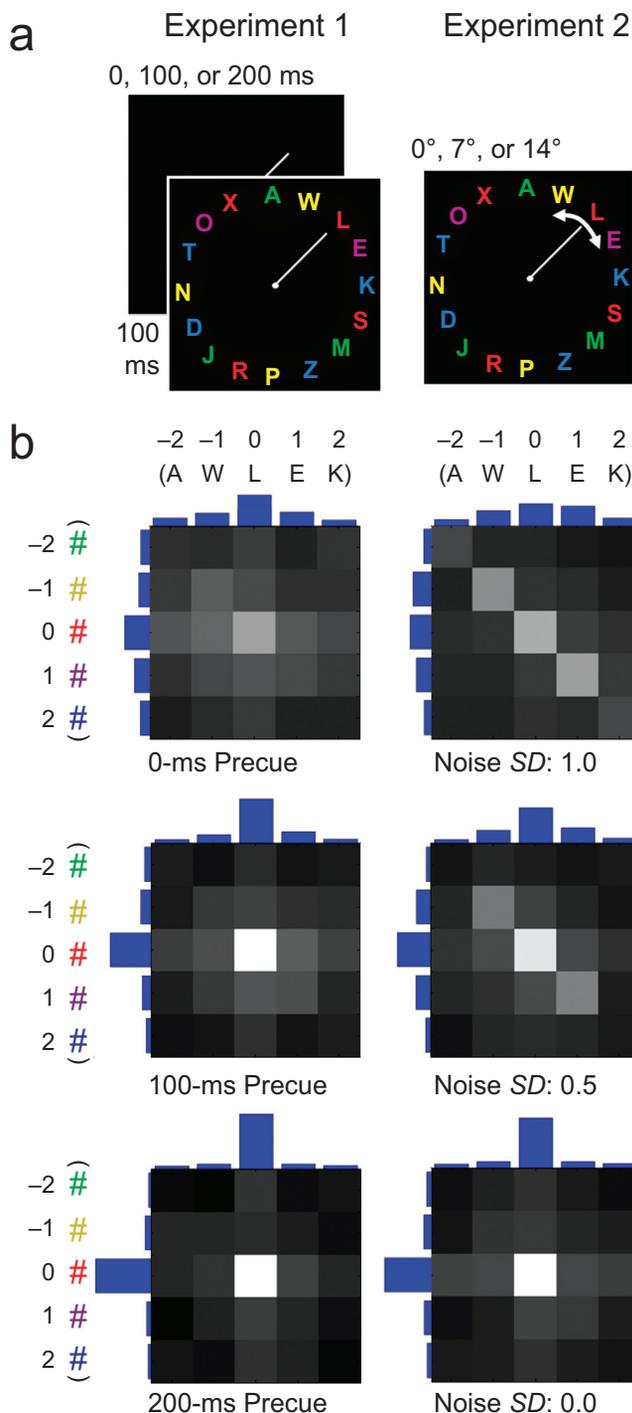
a

Experiment 1

0, 100, or 200 ms

100 ms

Experiment 2

0°, 7°, or 14°

b

0-ms Precue

100-ms Precue

200-ms Precue

Noise *SD*: 1.0

Noise *SD*: 0.5

Noise *SD*: 0.0

Fig. 1. Example trial sequences and results for Experiments 1 (left column) and 2 (right column). On each trial (a), an array of letters was presented in a ring around fixation for 100 ms. The target was cued by a white line. In Experiment 1, spatial uncertainty was varied by manipulating the time between onset of the cue and onset of the letter array (precue time). In Experiment 2, external noise was varied by perturbing the angular direction of the cue with Gaussian noise. The joint frequency histograms (b) denote the report frequency of each of the possible conjunctions in the three conditions of each experiment. Each row corresponds to the position of the reported color (from –2 to +2 relative to the target), and each column refers to a different letter identity (from –2 to +2 relative to the target); for illustration, the labels in parentheses refer to the specific example in (a). Increasing lightness reflects greater frequency of report. The marginal frequency histograms, shown in blue, denote the frequency of report of each feature separately.

eccentricity. The target was cued by a white line extending from the center of the display (~4° in length).

In the *spatial uncertainty* condition (Experiment 1), we manipulated the information available about the cue location (and, thus, the precision of attention) by accurately cuing the target location at a variable precue interval before the onset of the letter array. The time between onset of the cue and onset of the letter display (precue time) was 0, 100, or 200 ms. These values were chosen to discourage saccades to the target loca-tion.[1] These three precue conditions were randomly intermixed within a block. After the cue, the stimulus array was presented for 100 ms. Shorter precue intervals provided less information about the cue direction, thus decreasing spatial precision in esti-mated target locations. Our key question was whether the con-junctions of features would reveal that this imprecision was best described as arising from uncertainty or from internal noise.

The *spatial noise* condition (Experiment 2) was designed to illustrate the effects of shared noise on a given trial. For this condition, precue time was fixed at the longest duration (200 ms), but we added spatial noise to the location indicated by the cue. The noise was Gaussian, with a standard deviation equal to 1.0, 0.5, or 0 times the spacing of the letters (13.8, 6.9, and 0° of arc, respectively). These noise magnitudes approxi-mated the standard deviation (in items) of responses in Experi-ment 1. On average across trials, the cue pointed to the correct target, but on any one trial, it could point slightly off-target. External noise in the cued location simulated the possible con-tribution of internal noise in attended locations to feature intrusions in Experiment 1. We expected that position intru-sions would be correlated across features in this condition because, by design, color and identity reports shared a com-mon source of error: the external noise in cue position. Thus, this manipulation tested whether we could detect a correlation between feature intrusions when we knew there should be such a correlation.

Procedure. Participants completed five blocks of 60 trials each. Each trial began with a fixation cross for 500 ms. Next, the cue line appeared, and then the target display appeared after a variable (Experiment 1) or fixed (Experiment 2) inter-val. Participants used the number keys on a computer key-board to indicate the identity of the target letter, choosing from among five options; they chose from among five options in a separate display to report the target's color. Subjects were awarded 1 point for each feature reported correctly, so that the score for each trial could be 0, 1, or 2. The target location, color-letter pairings, position of each colored letter in the dis-play, and report order (color or identity first) were randomly chosen on each trial.

Results

Because the colors and letter identities of the target and four surrounding items were unique, we could identify the spatial position, relative to the cued item, corresponding to each reported feature. We used this information to construct the

joint distribution of color and letter reports—how often each of the 25 (5 letters × 5 colors) logically possible conjunctions was reported.

We quantified the spatial error of a given feature report by its spatial deviation: The deviation was 0 for the feature of the target, +1 for the feature of the item next to the target, and +2 for the feature of the item two positions away from the target. We (arbitrarily) labeled clockwise deviations as positive, and counterclockwise deviations as negative. Thus, in the example shown in Figure 1a, a report of "yellow" would be a –1 color intrusion, a report of "W" would be a –1 identity intrusion, and a report of "K" would be a +2 identity intrusion. The exact scoring is unimportant; what is critical is that we were able to calculate the magnitude and direction of spatial-position intrusions in both color and letter reports.

The variance of the spatial deviations describes the imprecision of feature reports. If this imprecision arose from independent sampling of color and letter identity given some spatial uncertainty about the location of the target, there would be no correlation between color and letter intrusions. In contrast, if this imprecision reflected internal noise in the estimated target location, then this noise would contribute to both color and letter errors, resulting in a correlation in their spatial intrusions. The covariance of deviations in color report and deviations in letter-identity report is a direct measure of the independence of feature intrusions. We measured the correlation of intrusions via their covariance (an unnormalized measure of the correlation[2]) because the units of covariance and variance were directly comparable. With this measure, we were able to test directly whether a spatial intrusion of color predicted an intrusion of letter identity and vice versa. A systematic relationship between errors on the two features would be detected as nonzero covariance in the joint distribution of errors.

Figure 1b (left column) shows the joint report distributions for the different precue durations (0 ms, 100 ms, and 200 ms) in Experiment 1. Participants reported the correct conjunction (central item) more often as the precuing time increased. This pattern demonstrates that the variation in precue duration successfully manipulated the spatial precision of attentional selection. The spatial variance of intrusions was lower with longer precue durations (Fig. 2): The change in variance as a function of precue duration in seconds had a slope between –6.5 and –3.5 (95% confidence interval, or CI) for letter reports and between –4.6 and –1.9 for color reports ($R^2 = .62$ and .47, respectively). Thus, with a longer cue exposure—and therefore more information about the cue's spatial direction—inference about which item was cued became more precise (or, alternatively, a narrower spatial window around the target was selected), increasing accuracy and precision of both color and identity reports.

Critically, despite the large changes in intrusion variance across precue conditions, precuing never increased the covariance of the feature intrusions: It remained at zero for all conditions (Fig. 2); 95% CIs were [–0.1, 0.19] for the 0-ms precue,
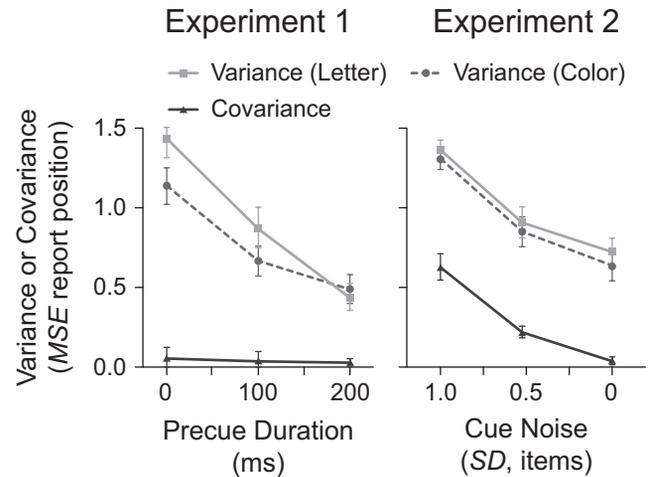

Fig. 2. Variance and covariance of the positions of feature intrusions (see the text) for the joint report distributions as a function of precue duration (Experiment 1; left) and external noise (Experiment 2; right).

[–0.08, 0.14] for the 100-ms precue, and [–0.03, 0.06] for the 200-ms precue. None of the precue conditions induced dependence between color and letter intrusions; feature intrusions were statistically equivalent to independent, identically distributed samples drawn from a probability distribution over space.

Previous work investigating the independence of feature reports suffered from an inability to distinguish whether visual selection or memory was the source of unbinding (and independence) in feature errors (Tsal, 1989). In Experiment 1, this was less of a problem, because we looked at spatial correlation rather than simply accuracy. Nonetheless, we further demonstrated the source of errors by contrasting the effects of spatial uncertainty about the cue location with the effects of external noise in the cue location.

In Experiment 2, all trials included the easiest precue (200 ms), but we perturbed the spatial position of the cue by adding to it external noise that was matched to the variability observed in Experiment 1; thus, the cue did not always point to the correct target (see Method; Fig. 1a). Effectively, this manipulation made the cue a less accurate indicator of target location. Because the error in the cue position would affect both color and letter reports, this external noise was expected to introduce a correlation between the feature reports. Thus, this experiment was designed to verify that our method could detect a correlation when the two features did in fact share a common source of error.

The results of Experiment 2 are shown in Figures 1b (right column) and 2. As expected, adding noise to the angular position of the cue decreased accuracy (i.e., the spatial variance of feature intrusions increased); the 95% CIs on the slope of this increase were [0.42, 0.89] for color reports and [0.36, 0.89] for identity reports ($R^2 = .48$ and .41, respectively). More important, cue noise added correlated variance: The covariance also increased (Fig. 2). The 95% CIs for this increase were as follows: [–0.03, 0.1] for 0 noise; [0.17, 0.28] for 7° noise; and

[0.46, 0.78] for 14° noise. When subjects made errors, the features they reported were correlated, and such correlations were detectable using our methodology. Thus, the independence we observed in Experiment 1 was not due to unbinding in memory or limited statistical power, but rather was due to independent sampling of features given spatial uncertainty in attentional selection.

Experiments 3 and 4: Binding in Time

Although binding in space is the canonical form of feature binding, illusory conjunctions and misbindings occur in time as well (Botella, Arend, & Suero, 2004; Botella, Suero, & Barriopedro, 2001). If random sampling of features is a general mechanism of visual attention (Vul, Hanus, & Kanwisher, 2009), our results should be replicable in the temporal domain. To test whether they were, we rearranged the 26 colored letters in a trial into a rapid serial visual presentation (RSVP) at fixation. One letter was cued by an annulus that appeared simultaneously with the letter (Fig. 3a). We manipulated the temporal uncertainty (precision of attentional selection in time) by varying the presentation rate (Experiment 3). Again, we contrasted this manipulation of uncertainty with a manipulation of external noise (Experiment 4).

Method

Participants. Participants were drawn from the Macquarie University subject pool and received course credit or were paid. There were 14 participants (10 female, 4 male; ages 18–45 years) in Experiment 3 (temporal uncertainty) and 12 participants (10 female, 2 male; ages 18–45 years) in Experiment 4 (temporal noise).

Materials. A rapid stream of capitalized Courier-font colored letters was presented at fixation, with a white ring cuing the target. Given the resolution (1024 × 768 pixels), monitor (Dell P992), and viewing distance (~57 cm), each letter subtended approximately 2.9° of visual angle.

In the *temporal uncertainty* condition (Experiment 3), the presentation rate was 13.3, 10, or 6.7 items/s; the item and interitem-blank durations were 45 and 30 ms, respectively, in the 13.3-items/s condition, 60 and 40 ms in the 10-items/s condition, and 90 and 60 ms in the 6.7-items/s condition. Trials in these three conditions were randomly intermingled. Cues appeared concurrently with targets.

In the *external noise* condition (Experiment 4), the presentation rate was fixed at 6.7 items/s, but we added temporal noise to the cue onset. The noise was Gaussian, with a standard deviation equal to 1.2, 0.8, or 0 times the item presentation time (180, 120, or 0 ms). Thus, in the nonzero-noise conditions, the cue onset could be slightly earlier or later than the target item's onset (and could even occur during items preceding or following the target).
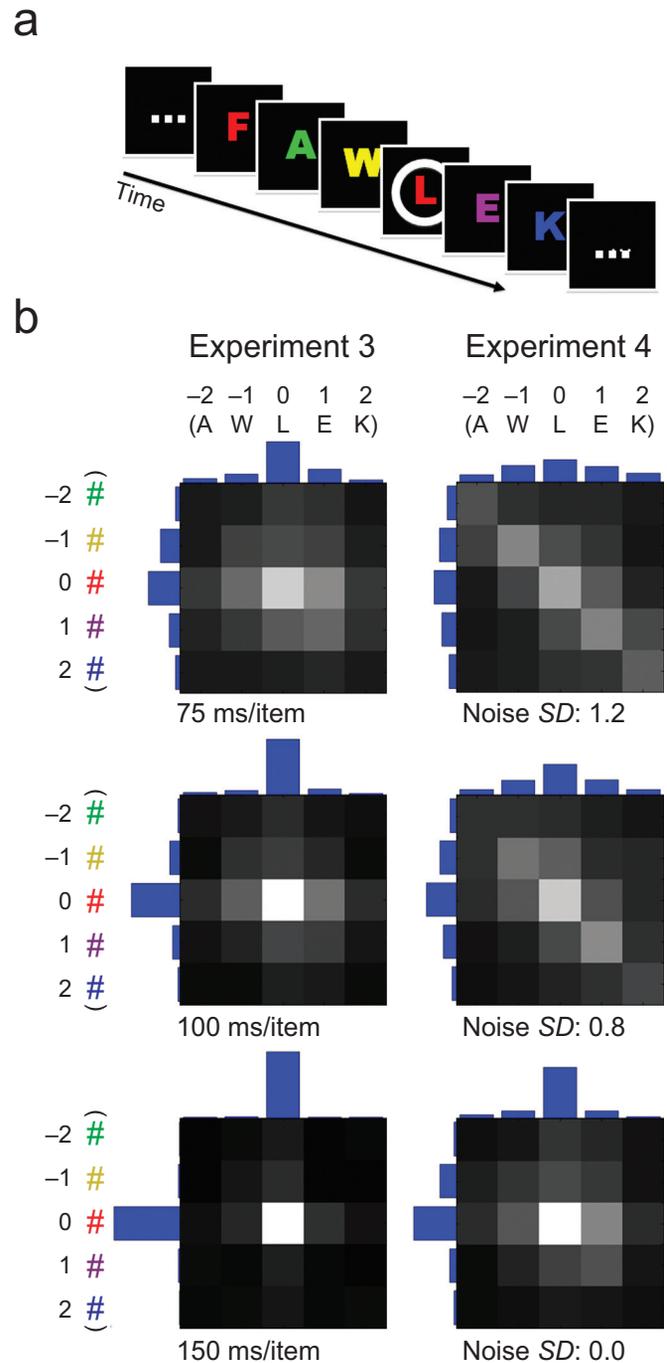

Fig. 3. Illustration of the trial sequence and results for Experiments 3 (left column) and 4 (right column). As illustrated in (a), letters were presented in a rapid sequence at fixation. The target was cued for report by a white annulus. In Experiment 3, temporal uncertainty in cuing was varied by manipulating the rate at which the items appeared. In Experiment 4, external noise was varied by perturbing the onset time of the cue with Gaussian noise. The joint frequency histograms (b) denote the frequency with which every possible conjunction was reported for the three conditions in each experiment. Each row corresponds to the position of the reported color (from −2 to +2 relative to the target), and each column refers to a different letter identity (from −2 to +2 relative to the target); for illustration, the labels in parentheses refer to the specific example in (a). Increasing lightness reflects greater frequency of report. The marginal frequency histograms, shown in blue, denote the frequency of report of each feature separately.

Procedure. Participants completed five blocks of 78 trials. As in the spatial experiments, subjects were awarded points for correctly reporting the color and letter identity of the target on each trial. The target position (Item 6 to Item 20, out of 26), letter order, color order, report order, and presentation rate (Experiment 3) or cue noise (Experiment 4) were all randomly chosen on each trial.

Results

For binding in time, just as in space, increasing temporal uncertainty of selection—in this case, by accelerating the RSVP stream—decreased accuracy. This is evident in the joint and marginal report histograms (Fig. 3b) and can be quantified by the variance of the temporal-position deviations of the reported features (Fig. 4). Variance of temporal intrusions increased with RSVP rate; the 95% CIs for the slope of this increase were [–10.2, –5.0] for color reports and [–12.0, –8.1] for letter reports (R^2 = .48 and .73, respectively). Just as we found with our manipulation of spatial uncertainty, manipulation of temporal uncertainty had virtually no effect on the dependence of feature intrusions. The covariance of intrusion positions did not correlate with RSVP rate; the 95% CI on the slope was [–1.1, 0.1] (R^2 = .076). Although neither the fastest nor the slowest RSVP rate showed a covariance significantly greater than zero (95% CIs of [–0.01, 0.09] and [–0.01, 0.02], respectively), the medium rate had a significantly nonzero covariance (95% CI: [0.03, 0.07]). However, compared with the overall variance of temporal intrusions, the magnitude of this nonzero covariance was negligible.

Figure 3b presents the joint and marginal report histograms for the three noise conditions in Experiment 4. In contrast to temporal uncertainty, the addition of external noise to the temporal position of the cue increased both the variance of letter and color reports and the covariance of feature intrusions

(Fig. 4). The 95% CI on the slope was [0.61, 0.89] for color reports and [0.54, 0.8] for letter reports (R^2 = .78 and .76, respectively). The 95% CI on the slope for covariance was [0.52, 0.72] (R^2 = .83).

The trend toward some nonzero covariance in the case of temporal uncertainty raises the possibility that there may have been some information about feature conjunctions in one of our attention conditions. We hesitate to draw this conclusion, however, because the effect was so small compared with the fluctuations in intrusion variance: The correlation between color and letter intrusions accounted for an average of 8% of the position errors, and the magnitude of the covariance (unlike variance) did not change with temporal uncertainty. In contrast, when we manipulated temporal noise, covariance changed with variance, and accounted for 43% of the variability in letter and color intrusions. Thus, we conclude that in time, just as in space, feature reports are well described as statistically independent samples from a spatiotemporal probability distribution.

Discussion

Attentional feature binding appears fundamental to conscious visual experience, allowing observers to effortlessly perceive objects as cohesive structures of different features. We examined how selective attention achieves this critical process by asking what determines the perceived features of an object when attention fails to create the veridical conjunction. Our results show that two features perceived as a conjunction are statistically equivalent to two independent samples from a probability distribution (attentionally selected region) in both space and time. Accurate binding of features is therefore not a special mechanism or action by selective attention, but merely the limiting case when the attentionally selected region is narrow enough to encompass only one object. According to our account, this happens when there is sufficient information about the spatiotemporal location of the attended object.

There are a few alternative explanations of the source of feature intrusions in binding. First, there may be internal noise in estimating the spatiotemporal location of an object. That is, noise in the visual system may result in selection of an incorrect location. Our data rule out this account, however, because the letter and the color reports would have shared this noise, and therefore their intrusions would have been correlated. Second, independence between different feature intrusions may arise from independent spatiotemporal noise for each feature. One way to test for such feature noise is to get participants to make multiple guesses about a single feature (e.g., guess the cued letter's identity and then make a second, different, guess about the cued letter's identity); if a given feature contains noise, then errors across two guesses about one feature should be correlated. Recent data, however, show that multiple guesses about one feature also contain independent error (Vul, Hanus, & Kanwisher, 2009). Thus, the intrusions in our present results seem unlikely to have arisen from internal noise on individual feature dimensions.
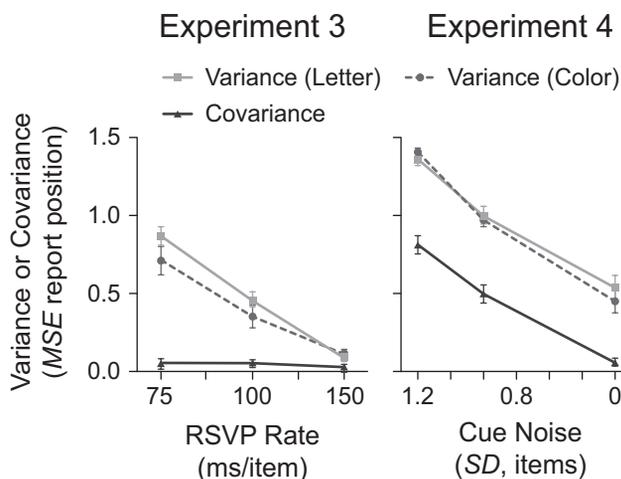

Fig. 4. Variance and covariance of the positions of the feature intrusions (see the text) for the joint report distributions as a function of rapid serial visual presentation (RSVP) rate (Experiment 3; left) and external noise (Experiment 4; right).

Our interpretation of the current results is that the rapid presentation of a spatial or temporal cue provides insufficient information about the spatial or temporal position of the cued object. Participants are left with some uncertainty about the location of the relevant item, and the "window of attentional selection" corresponds to a probability distribution over space and time, describing the inherent uncertainty of the task. This spatiotemporal uncertainty yields a probability distribution over features likely to have been present in that location. Subjects then independently sample features from this probability distribution. Thus, the color and letter responses in our experiments were sampled from probability distributions that encompassed the likely target as well as the surrounding items. Crucially, the two features were sampled independently, as demonstrated by a lack of correlation in their errors. According to this account, both veridical and illusory binding arise from the way visual attention copes with uncertainty: approximation through sampling (Vul, Goodman, Griffiths, & Tenenbaum, 2009; Vul, Hanus, & Kanwisher, 2009).

These results connect binding to a growing literature that suggests that, in general, the human mind implements complex probabilistic computations via sampling, which results in responses that appear to be probability-matched to beliefs (Goodman, Tenenbaum, Feldman, & Griffiths, 2008; Herrnstein, 1961; Sanborn & Griffiths, 2008; Vul, Goodman, et al., 2009; Vul, Hanus, & Kanwisher, 2009; Vul & Pashler, 2008). Our data suggest that visual attention acts as a sampling process to select visual features for conscious perception, and that there is no additional "binding" process: Veridical binding is just the limiting case of this sampling process, when the spatiotemporal window from which features are independently sampled is narrow enough to contain only one object.

Acknowledgments

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

Notes

1. On each trial, we asked subjects if they had moved their eyes (to remind them to fixate), and we discarded trials on which they reported having done so. Eye movements are not a major concern because saccades to the target location would induce (rather than mitigate) a correlation in feature reports.

2. The relation between correlation and covariation can be expressed as follows: $r_{xy} = \sigma_{xy}/(\sigma_x \sigma_y)$, where r_{xy} is the correlation, σ_{xy} is the covariance, and σ_x and σ_y are the marginal standard deviations of x and y.

References

Alvarez, G.A., & Oliva, A. (2008). The representation of simple ensemble features outside the focus of attention. *Psychological Science*, *19*, 392–398.

Alvarez, G.A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences, USA*, *106*, 7345–7350.

Ashby, F.G., Prinzmetal, W., Ivry, R., & Maddox, W.T. (1996). A formal theory of feature binding in object perception. *Psychological Review*, *103*, 165–192.

Botella, J., Arend, I., & Suero, M. (2004). Illusory conjunctions in the time domain and the resulting time-course of the attentional blink. *Spanish Journal of Psychology*, *7*, 63–68.

Botella, J., Suero, M., & Barriopedro, M.I. (2001). A model of the formation of illusory conjunctions in the time domain. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 1452–1467.

Bundasen, C., Kyllingsbaek, S., & Larsen, A. (2003). Independent encoding of colors and shapes from two stimuli. *Psychonomic Bulletin & Review*, *10*, 474–479.

Chong, S.C., & Treisman, A. (2005). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, *67*, 1–13.

Goodman, N., Tenenbaum, J., Feldman, J., & Griffiths, T. (2008). A rational analysis of rule-based concept learning. *Cognitive Science*, *32*, 108–154.

Herrnstein, R. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*, 606.

Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review*, *114*, 599–631.

Huang, L., Treisman, A., & Pashler, H. (2007). Characterizing the limits of human visual awareness. *Science*, *317*, 823–825.

Humphrey, G.K., & Goodale, M.A. (1998). Probing unconscious visual processing with the McCollough effect. *Consciousness and Cognition*, *7*, 494–519.

Isenberg, L., Nissen, M.J., & Marchak, L.C. (1990). Attentional processing and the independence of color and orientation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 869–878.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement and depth: Anatomy, physiology and perception. *Science*, *240*, 740–749.

Monheit, M.A., & Johnston, J.C. (1994). Spatial attention to arrays of multidimensional objects. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 691–708.

Nissen, M.J. (1985). Accessing features and objects: Is location special? In M.I. Posner & O.S.M. Marin (Eds.), *Attention and performance XI* (pp. 205–219). Hillsdale, NJ: Erlbaum.

Robertson, L.C. (2003). Binding, spatial attention and perceptual awareness. *Nature Reviews Neuroscience*, *4*, 93–102.

Sanborn, A., & Griffiths, T. (2008). Markov chain Monte Carlo with people. In J.C. Platt, D. Koller, Y. Singer, & S. Roweis (Eds.), *Advances in neural information processing systems 20* (pp. 1265–1272). Cambridge, MA: MIT Press.

Sincich, L.C., & Horton, J.C. (2005). The circuitry of V1 and V2: Integration of color, form, and motion. *Annual Review of Neuroscience*, *28*, 303–326.

Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition*, *14*, 411–443.

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

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*, 107–141.

Tsal, Y. (1989). Do illusory conjunctions support the feature integration theory? A critical review of theory and findings. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 394–400.

Vul, E., Goodman, N.D., Griffiths, T.L., & Tenenbaum, J.B. (2009). One and done? Optimal decisions from very few samples. In *Proceedings of the 31st Annual Meeting of the Cognitive Science Society, Amsterdam, Netherlands* (pp. 148–153). Hillsdale, NJ: Erlbaum.

Vul, E., Hanus, D., & Kanwisher, N. (2009). Attention as inference: Selection is probabilistic; responses are all-or-none samples. *Journal of Experimental Psychology: General*, *4*, 546–560.

Vul, E., & MacLeod, D.I. (2006). Contingent aftereffects distinguish conscious and preconscious color processing. *Nature Neuroscience*, *9*, 873–874.

Vul, E., & Pashler, H. (2008). Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science*, *19*, 645–647.

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