Feature binding across different visual dimensions

Robert Ward · Isabel Arend

Published online: 12 June 2012
© Psychonomic Society, Inc. 2012

Abstract The human brain represents different kinds of visual feature dimensions in different ways. For example, surface features exhibit some properties that are very different from contour features, and some feature dimensions may be represented more extensively in either the dorsal or the ventral visual stream. Given such differences, we investigated feature binding across different feature dimensions and whether some feature dimensions might be more easily bound together than others. In [Experiment 1](#), we looked at cross-dimension bindings for all combinations of color, orientation, and shape dimensions, while at the same time controlling for feature discriminability. Rates of correct binding, illusory conjunctions, and feature errors were equivalent in all cases. There was no bias so that some feature dimensions were easier to combine than others. In [Experiment 2](#), we manipulated the difficulty of feature discrimination for the key, target-defining feature and the report feature. Rates of binding errors increased with difficulty of the key feature, but not with that of the report feature. The accuracy of feature discrimination could be dissociated from the accuracy of binding the feature to an object. Across both experiments, the accuracy of feature binding was independent of specific feature dimensions or perceptibility. These findings are discussed in relation to both feature integration and multiple-stage accounts of visual feature integration.

Keywords Illusory conjunctions · Feature integration · Spatial coding · Task difficulty

Feature binding across visual dimensions

Visual processing is distributed in parallel across many specialized brain areas. Early findings identified brain areas specialized for features such as color, form, and motion (e.g., Livingstone & Hubel, 1988), and more recently, specializations for complex stimuli, including faces, scenes (Epstein, Harris, Stanley, & Kanwisher, 1999), and bodies (Downing, Jiang, Shuman, & Kanwisher, 2001), have been identified. This distribution of processing raises the problem of coordination, so that different visual processors provide analysis of the same object at the same time (Duncan, 1993). Logically, such coordination is essential for many fundamental visual operations; for example, to identify a target in a crowded field on the basis of its shape and then report its color, form and color processing must be linked to the same object. Behaviorally, coordination is demonstrated by the *object-based* nature of visual attention. For example, under limited exposures, two features of an object can be reported as accurately as one feature; however, reporting one feature from two different objects leads to robust performance costs (Duncan, 1984).

A binding mechanism that integrates outputs from multiple visual processors was formulated in the influential feature integration theory (FIT; Treisman & Gelade, 1980). According to FIT, *illusory conjunctions* are miscombinations of visual features that, in the absence of focal attention, are registered with little or no assignment to spatial location (Treisman & Schmidt, 1982). Focal attention to location acts to bind the output of feature processors for that location. In

R. Ward (✉) · I. Arend
Wolfson Centre for Clinical and Cognitive Neuroscience,
Bangor University,
Bangor LL57 2AS, UK
e-mail: r.ward@bangor.ac.uk

I. Arend
Department of Psychology, Ben-Gurion University of the Negev,
Beer Sheva, Israel

this way, the theory illustrates the link between feature binding and the coordination of multiple visual areas in an object-based, or at least location-based, manner. FIT has led to subsequent work demonstrating that space is an important medium for establishing feature bindings. Features that appear close together are more likely to be conjoined than more distant features (e.g., Cohen & Ivry, 1989; Wolford & Shum, 1980). With long exposures (up to 1.5 s) and low attentional load, illusory conjunctions can be obtained in the periphery (Prinzmetal, Henderson, & Ivry, 1995), but not at the fovea, where spatial coding is more precise (Treisman & Schmidt, 1982). Despite modifications over the years, FIT has maintained the central claim that spatial attention is the main mechanism for the integration of independent feature codes (Quinlan, 2003). In this way, FIT proposes a unitary mechanism for feature binding.

A recent alternative to FIT also argues for a central role of spatial coding but, rather than a unitary binding mechanism, proposes a model of multiple, dissociable binding processes (Humphreys, 2001; see also Humphreys, Cinel, Wolfe, Olson, & Klempen, 2000; Humphreys, Hodsoll, & Riddoch, 2009). Whereas FIT constitutes a conceptual descriptive account on how features are integrated, the multistage account goes beyond the descriptive level to propose that different binding mechanisms have a specific neuroanatomical basis. According to the multistage account, visual features are registered, and visual forms are grouped into objects. Grouping between form elements is meant to occur largely preattentively, on the basis of both bottom-up grouping and top-down reentrant processing of stored object representations, from the ventral stream. Form elements are still subject to miscombination, since large receptive fields within the ventral stream can fire to the attributes of multiple objects in the visual field. Color information is also processed in early stages but only later integrated with shape. Spatial attention acts subsequently to stabilize these early conjunctions, possibly by retuning the receptive fields along ventral areas. In addition, it is at this stage that bindings are made across different feature dimensions, such as color and form. Similar to the central role of spatial attention in FIT, activation from spatiotopic parietal areas feeds down to feature locations, to minimize binding errors. The multistage account assumes that top-down activation from the object recognition system, derived from early binding, will “push” the parietal system so that the coactivated locations become attended. Therefore, even while dorsal and ventral visual streams demonstrate specialization of function, their activity is coordinated to promote feature binding.

Both FIT and the multistage model illustrate the link between feature binding and the coordination of multiple visual areas in an object-based manner. Both accounts also agree on the importance of spatial processing as a basis for feature binding. However, FIT considers a unitary binding

mechanism (possibly coordinated by the parietal cortex) that operates irrespective of the nature of the feature dimension (i.e., orientation, shape, or color). In contrast, the multistage account considers that binding involves different processes, which are dependent on the nature of the feature dimension. For example, binding involving shape and color would be acquired through the dorsal pathway (parietal mechanisms), whereas within-form binding would be acquired through the ventral pathway (occipital-temporal areas).

The multistage model is developed from results showing that binding across different feature dimensions is dissociable from binding within the *form* dimension (Humphreys et al., 2000; Humphreys et al., 2009). An example of binding within the form dimension might be integrating a vertical and horizontal segment to make either a “T” or an “L.” Two studies argue for the distinction between form binding and cross-dimension binding. Humphreys et al. (2000) examined a patient with bilateral damage to the parietal cortex and evidence of Balint’s syndrome. The patient showed evidence of intact form binding but was significantly impaired at binding across visual feature dimensions, such as color and shape. Further evidence for a distinction between cross-dimension conjunctions and conjunctions within the form dimension was found by Humphreys et al. (2009). Humphreys et al. (2009) found that parietal patients had specific difficulty with cross-dimension feature conjunctions, but not form conjunctions. The parietal cortex therefore appears to play a critical role in cross-dimension binding. Recently, this account has received support from neuroimaging findings (Wei, Muller, Pollman, & Zhou, 2011) showing dorsal activity for cross-dimension search and occipital-temporal activity for within-dimension search.

The dissociation between within-form and cross-dimension bindings demonstrated in the multistage model raises the question of whether there may be mechanisms specific for different cross-dimension bindings. That is, are some feature conjunctions, such as size and shape, color and motion, and so forth, more easily integrated than others? The theoretical picture is ambiguous. The FIT and multistage model both propose a mission-critical but general spatial attention process underlying binding efficiency across different feature dimensions. Binding errors might then be expected to be equally likely for different combinations of feature dimensions, as long as spatial processing is not allocated to the appropriate locations. That is, even though the multistage model predicts dissociations of cross-dimension and within-form bindings, it is agnostic about different forms of cross-dimension binding.

However, there are some plausible reasons to predict that binding efficiency might not be the same for different cross-dimension bindings. For one, neural representations responding to specific feature conjunctions, such as receptive field tuning to both shape and color (Tanaka, Saito, Fukada, &

Moriya, 1991), might allow highly efficient feature integration. Second, different visual features may be coded at different spatial resolutions. Given the importance of accurate spatial coding for binding (Ward, Danziger, Owen, & Rafal, 2002), if some features are represented by less precise spatial codes, they might be more easily miscombined with features at nearby locations. For example, there seems to be a difference between surface features, such as color, and contour features, such as shape. This possibility is illustrated by the filling-in phenomena, such as neon color spreading. Neon color spreading is a filling-in illusion in which a part of an object's surface is perceived as transparent even though the object itself has no transparent physical surface (Sasaki & Watanabe, 2004). Finally, given the well-known segregation of dorsal and ventral streams, it is not implausible that binding might be differently efficient for features represented *within*, as compared with *between*, streams. For example, conjunctions of color and motion might be difficult to integrate to the extent that ventral and dorsal streams are functionally separated (Milner & Goodale, 2008). At present, then, the theoretical picture seems to allow either similarity or differences in the processing of different cross-dimension conjunctions.

The empirical picture is also ambiguous. Following the methods of Treisman and Schmidt (1982), the majority of studies on cross-dimension binding seem to have focused on conjunctions of color and form, and only a few studies have systematically assessed a variety of cross-dimensional bindings. The available studies are generally related to the issue of identifying qualitative differences in visual search for single features and feature conjunctions. For example, Treisman and Sato (1990) found that all the feature conjunctions tested (numerous combinations of size, motion, color, and orientation) produced search slopes significantly different from the shallow slopes of feature-based search. There were differences in search slopes between conditions, but they were irrelevant to the main purpose of Treisman and Sato's study, and little could be made of these, due to differences in single-feature conditions.

Some single-case studies of neurological patients with Balint's syndrome have examined a variety of cross-dimension bindings. For example, Grabowecky, Robertson, and Treisman (1993) tested efficiency of a patient with Balint's syndrome in binding different combinations of shape, color, and motion. The patient produced illusory conjunctions irrespective of feature dimension, although motion produced more intrusions than did shape and color. However, the different feature dimensions were not calibrated in this case, and so quantitative comparisons across conjunction types is impractical to assess.

Humphreys et al.'s (2009) study is unique, to our knowledge, in comparing two different kinds of cross-dimension binding that have been equated for difficulty. Humphreys et al. (2009, Experiment 3) titrated exposure duration on the small

display search task so that performance in the color-size and form-size conjunction tasks was equal. Note that this duration adjustment also considered performance on the single-feature search tasks. Indistinguishable search slopes for color-size and form-size conjunctions were then found when this exposure duration was used to test search displays of variable size. Of course, this is a very small database of potential conjunctions, but these results would suggest that different kinds of cross-dimensional conjunctions are processed in similar ways, unaffected by specific feature dimensions.

Here, we systematically examine the efficiency of binding different feature dimensions. We measure illusory conjunctions, or binding errors, for different cross-dimension conjunctions, while controlling for the rate of nonbinding, perceptual errors. Only by controlling feature discriminability in this way can we determine whether some feature dimensions are more easily bound together than others.

Experiment 1: Binding across shape, orientation and color dimensions

Our method was generalized from previous studies of visual feature binding (Cohen & Rafal, 1991; Ward et al., 2002). On each trial, a target and a distractor were briefly presented to the left or right field and then masked. We adopt Botella, Barriopedro, and Suero's (2001) terminology, and so we refer to the feature that defines the target as the *key feature* (KF). For example, given that the participant's task is to report the color (with different feature values, such as red, green, and blue) of the only uppercase word presented among other lowercase words, the KF refers to the value uppercase, since it differentiates the target from the distractors. The feature that the observer is asked to report is called the *report feature* (RF). In this case, the RF can have three different values: red, green, and blue.

In the present study, the target and distractor were manipulated along the KF and RF dimensions. In all tasks, there were always three possible values for the KF and RF (see the [Method](#) section). Taking as an example our task in which the KF dimension was color and the RF dimension was shape, the participants' task was then to indicate which one of two possible values of the KF was presented on each trial (i.e., red or green; the third value was always associated to the distractor) and to indicate the shape of the target (i.e., letters S, E, C). Therefore, if the KF was red or green, the distractor was always blue. Regarding the RF trials, two out of three possible features were always present, one as the target and the other as the distractor.

We systematically assessed binding efficiency for each of the six combinations of KF and RF, made possible by cross-dimension pairings of shape, color, and orientation. We equated the perceptual difficulty of both the KF and RF on

each task, so that we could directly compare binding accuracy on the different pairings.

We take one of the displays used in the present study to exemplify the three possible outcomes involving the target and the distractor features (see Fig. 1). Because there were three values for the RF, there were three possible outcomes on any given trials for RF identification. First, the RF of the target could be identified correctly (reporting red S in Fig. 1). Second, the RF of the distractor could be mistakenly reported as belonging to the target (red E). This is an example of a *binding error*, or illusory conjunction, in which a feature from one object is mistakenly combined with the features of a different object. Finally, the RF reported might not be present in the display (red C). This is a *feature error*, indicating a problem in feature perception, not in binding.

A staircasing procedure was used to keep the rate of feature errors roughly at midway between perfect and chance identification (as indicated by 0 % and 33 % feature error rates, respectively). A fundamental point is that the RF staircasing controlled the feature error rate (i.e., the perceptual errors on the report dimension) but had no direct influence on the ratio of correct responses to binding errors. To illustrate, consider two different binding processes: perfect binding, which conjoins together features only from the same object, and chance binding, which combines features from different objects at random. Assume further that the rate of feature errors in both cases is staircased to 20 %. At one extreme, with perfect binding, the remaining 80 % of trials would all be correct responses. But as binding efficiency drops, the same 80 % of trials will include progressively more illusory conjunctions. Eventually, the other extreme would be reached, of chance levels of binding. At this extreme, of chance binding, the 80 % of trials without feature errors would be evenly split between correct responses and binding errors. The use of RF staircasing therefore ensures that perceptual difficulty of the RF is equivalent across tasks but allows any differences in binding efficiency to emerge.

Fig. 1 Example display in which orientation is the key feature, and shape is the report feature. Orientation distinguishes the target (oblique) from the distractor (horizontal). On this basis the correct shape report is “S”, a binding error would be “C” (the shape of the distractor), and a feature error would be “E” (the report feature value not present)

Method

Participants

There were 60 participants (39 female; mean age = 19 years), 10 in each task, receiving £5 or course credit for the 1-h experiment. This total excludes 14 participants for whom the staircasing procedure was ineffective. For 7 participants, the KF discrimination could not be made difficult enough (average performance of this group = 91 %; chance performance = 50 %; staircasing target = 75 %); for 7 others, the KF could not be made easy enough (average performance = 61 %). All participants reported normal or corrected-to-normal vision acuity.

Design

The dimensions of shape, color, and orientation could be either the KF (the two possible feature values that defined the target) or the RF (the three possible feature values to be reported). Pairwise combinations of KF and RF resulted in six experimental tasks: shape/color, color/shape, color/orientation, orientation/color, shape/orientation, and orientation/shape. Rather than analyze our findings on the basis of a single six-level factor, most of our analyses look separately at performance as a function of KF and of RF. The stimuli could appear in either the left or the right visual field, which was treated as a within-subjects factor, and task was treated as a between-subjects factor.

Tasks

For each dimension, there were three feature values: color (red, green, blue), orientation (left tilt, right tilt, horizontal), and shape (C, E, or S). When a dimension was used as the KF, one value was always fixed for the distractor, and the target was equally likely to be one of the other two values: For shape as the KF, the distractor was always a “C,” and the target either “E” or “S”; for color, the distractor was always blue, and the target either red or green; for orientation, the distractor was always horizontal, and the target tilted to either the left or the right. When the dimension was the RF, on each trial, one of the three values was randomly assigned to the target, one was assigned to the distractor, and one was not presented.

Procedure

Each trial began with a white fixation cross on an otherwise black screen. The participant pressed the space bar to initiate the trial. After a delay of 300 ms, the fixation was replaced by the stimulus display, containing a single central digit, the target, and the distractor. The digit varied randomly from 1

to 4 on each trial. The target and distractor always appeared at a horizontal eccentricity of 1.5° from center and were vertically separated by 0.4° from each other from a viewing distance of 55 cm. The target and distractor were centered within an area of $0.6^\circ \times 0.6^\circ$ of visual angle from a viewing distance of 55 cm. The stimulus display appeared for 100 ms. A mask display then appeared for 150 ms, composed of four white asterisks, one at each possible target and distractor location and subtending $0.5^\circ \times 0.5^\circ$ of visual angle.

Following the mask display, the participants pressed labeled keys to indicate response to the central digit (four possible values), the KF (two possible values), and the RF (three possible values). The instructions made clear that the digit task was primary. This central task was used to both ensure fixation and increase binding errors, by limiting attention to the periphery. Each experiment lasted for approximately 50 min, in which the participants performed 30 practice and 288 experimental trials.

A simple staircase procedure was used to equate the difficulty of KF and RF discriminations across tasks, and it was applied during the entire duration of the experiment. The staircase was set to increase the difficulty of the KF discrimination after five consecutive correct responses and to decrease difficulty after one incorrect response. The RF was staircased on the basis of the number of feature errors made. A feature error refers to the cases in which a feature that was not presented in the display was reported as the target feature. After eight consecutive, nonfeature error reports (i.e., when the report feature of the target was identified correctly), the RF discrimination was increased, and after a feature error, the RF discrimination was decreased in difficulty. In this way, KF staircasing controlled the rate of target identification errors, while RF staircasing controlled the rate of feature errors. In both cases, the rate of illusory conjunctions was free to vary with the task characteristics. Monte Carlo simulations (Mooney, 1997) showed that this procedure would result in an asymptotic KF error rate around 24% (chance performance would be 50 %) and a feature error rate of around 19 % (chance performance would be 33 %). As described earlier, we replaced participants for whom our staircasing did not prevent either very low or very high performance on either the KF or RF.

Results and discussion

Accuracy on the primary digit task was 88.5 %, and there were no main effects or interactions involving KF or RF dimensions, at an alpha level of .05. We were therefore satisfied that the primary digit task was performed at a high and equivalent level across tasks.

We next looked at the effects of our staircasing manipulations. Accuracy on the KF across all tasks averaged 77 % (or 23 % error rate, close to our Monte Carlo estimate of 24 %). As is shown in Fig. 2, the KF dimension—color, orientation, or shape—had little effect on KF accuracy, $F(1, 57) = 3.0, p = .09$. The other staircased parameter was tied to the feature error rate. The mean feature error rate across all tasks was 16 %, which was again close to our Monte Carlo estimate of 19 %. Again, as is shown in Fig. 2, the report feature dimensions—color, orientation, or shape—had little effect on feature error rate, $F(1, 57) = 1.11, p = .34$.

Of main importance is the rate of illusory conjunctions. We examined this issue in several ways, but the outcome was always the same: Cross-dimensional binding accuracy did not significantly vary as a function of feature dimensions. Illusory conjunctions occurred on 26 % of trials, and this rate was unaffected by the specific feature dimensions presented. There was no significant effect of KF dimension on illusory conjunction rate, $F(2, 57) = 0.138$, or of RF dimension, $F(2, 57) = 0.567$. A priori power analysis indicated that the 60 participants provided reasonable power, so that the power to detect an effect size of .5 was .932. Given a beta/alpha ratio of 1.0 (i.e., weighting the importance of Type I and II errors equally), compromise analysis suggested that the power to detect this effect size was .941.

Another way to assess any effect of cross-dimensional binding is to consider the possibly distinct effects of each of the six tasks—that is, as a six-level factor in an analysis of variance (ANOVA). No effect of the six tasks on illusory conjunction rate was found, $F(5, 54) = 0.254$. For this ANOVA, a compromise power analysis, given a beta/alpha

Fig. 2 Key (KF) and Report feature (RF) accuracy in Experiment 1 (error bars represent 99 % confidence interval). For the KF discrimination, chance was 50 %, and the staircase maintained an overall error rate of 23 % that did not differ between feature dimensions (color, orientation, and shape). For the RF discrimination, chance performance was 33 %, and the staircase maintained an average error rate of 16 %, that did not differ between dimensions

ratio of 1.0, suggested that the power to detect this effect size was .900. It therefore seems likely that, when the perceptual difficulty of the KF and RF is carefully controlled, there are no large differences in illusory conjunction rates for the different cross-dimensional bindings we examined.

Equivalent results are obtained for analyses of illusory conjunctions and for correct identifications. This is simply because total RF performance is the sum of correct identifications + illusory conjunctions + feature errors. We have already seen that illusory conjunctions and feature errors are highly similar across tasks, and therefore, so too must be correct RF identification. Correct RF identification was unaffected by KF dimension, $F(2, 57) = 0.137$, or RF dimension, $F(2, 57) = 0.086$. Figure 3 shows the breakdown of RF reports across tasks and response categories.

These similarities in illusory conjunction rate are not the result of floor or ceiling effects, in either feature identification or binding. Our staircasing ensured that perception of both relevant feature dimensions was approximately midway between floor and ceiling. If binding performance had been at ceiling, we would have expected no illusory conjunctions. If binding efficiency had been at chance, we would have expected roughly equal numbers of correct RF identifications and illusory conjunctions. Instead, we observed a ratio of about 25 % conjunction errors (or about 72 such errors in the 288 trials for a participant) to 58 % correct identifications (or about 167 of 288 trials).

However, it is the case that the observed 25 % rate of illusory conjunctions is an overestimate of true binding errors. The rate of illusory conjunctions reflects not just true binding errors, but also feature misperceptions. That is, blind guesses of the RF would make a response classified as an illusory

conjunction on one third of trials (since there were three possible values for the report feature). Because random guessing would produce equal numbers of feature errors and illusory conjunctions (as well as equal numbers of correct responses), and because the feature error rate is a reasonable estimate of uninformed guessing, a conservative approach would be to subtract the feature error rate from the illusory conjunction rate to estimate the true rate of binding errors. A similar approach was taken by Treisman and Schmidt (1982). If we estimate *binding error rate* for each participant as the difference between the illusory conjunction rate (26 %) and the feature error rate (16 %), binding errors occurred on about 10 % of the trials. Although this estimation process is interesting, it has no effect on our main conclusion, that the efficiency of cross-dimensional binding is similar across dimensions. Because we controlled the feature error rate across all tasks, any form of correcting for the effect of feature errors on estimates of binding efficiency makes no difference between tasks.

Experiment 2: Attention and binding

Experiment 1 systematically examined a variety of cross-dimension feature bindings and found equivalent performance in all cases. These results are consistent with a single mechanism for binding across visual-feature dimensions. These results raise the following question: What does affect binding difficulty? In our previous experiment, we followed Botella et al. (2001) in distinguishing between the KF, which differentiates the target from the distractor, and the RF. Experiment 1 tightly controlled the perceptual difficulty of both. In this experiment, we examined how difficulty in processing the KF and RF influences binding.

Fig. 3 Frequency of the three alternative responses for the report feature (RF), for each task (the combination of key feature and report feature). Error bars indicate 99 % confidence intervals

Method

The general method was similar to that of [Experiment 1](#). In this case, we used only a single task, with shape as the KF and color as the RF. For one group of participants, the staircasing of the KF was adjusted so that the discrimination was significantly more difficult than in [Experiment 1](#); for the other group, the RF discrimination was made more difficult (see the [Procedure](#) section).

Participants

There were 20 (12 female; mean age = 21 years, $SD = 3$) new participants from the Bangor participant panels, receiving either £5 or course credit for the 1-h experiment. Four of these participants were excluded because the staircasing was ineffective. All participants reported normal or corrected-to-normal vision acuity.

Design

A single between-subjects factor (*hard KF* or *hard RF*) indicated whether the KF or RF discrimination was made difficult relative to [Experiment 1](#).

Procedure

The procedure was identical to the shape/color task in [Experiment 1](#), with the exception of the staircasing procedure. In the hard KF condition, the difficulty of the KF discrimination was increased. In [Experiment 1](#), KF difficulty was increased after five correct KF discriminations. In the hard KF condition, this was changed to increase difficulty after three correct responses. A Monte Carlo simulation suggested that this would produce KF accuracy around 68 %. Similarly, in [Experiment 1](#), RF difficulty was increased only after eight nonfeature error responses (i.e., correct RF responses). In the hard RF conditions, this was changed to five correct responses. Otherwise, staircasing in both conditions was identical to that in [Experiment 1](#).

Results and discussion

Accuracy on the digit task was 89 %, very comparable to that in [Experiment 1](#). Accuracy on the primary task was unaffected by whether the hard discrimination was on the KF or RF, $F(1, 18) = 0.98$.

We next verified that our staircasing procedures had the desired effect. As was expected, KF accuracy was lower in the hard KF than in the hard RF condition (63 % and 77 %, respectively), $F(1, 18) = 54.6$, $p < .0001$. Also, just as was expected, feature errors were higher in the hard RF than in

the hard KF condition (21 % and 15 %, respectively), $F(1, 18) = 54.7$, $p < .0001$.

The question of main interest is whether binding efficiency differed in the hard KF and hard RF conditions. Binding was more difficult in the hard KF condition, evidenced by greater numbers of conjunction errors in the hard KF condition, than in the hard RF condition (33 % and 27 %, respectively), $F(1, 18) = 4.61$, $p = .046$.

However, this direct comparison is biased against finding a true difference in binding. As was discussed in [Experiment 1](#), the rate of illusory conjunctions reflects not just true binding errors, but also feature errors. In [Experiment 1](#), this was not crucial, because we controlled feature errors across tasks. However, in this experiment, we deliberately manipulated the feature error rate so that it was greater in the hard RF than in the hard KF condition. It is therefore interesting that it was still the hard KF condition that showed more illusory conjunctions. However, if we compare the binding error rate (illusory conjunction rate – feature error rate), the difference in binding errors for the hard KF over the hard RF condition becomes even greater (17.5 % and 7 %, respectively), $F(1, 18) = 17.85$, $p = .001$.

The simple and conservative approach suggested by Treisman and Schmidt (1982) is not the only way to correct estimates of binding errors for the influence of feature errors. More sophisticated models, such as those of Prinzmetal, Ivry, Beck, and Shimizu (2002) and Ashby, Prinzmetal, Ivry, and Maddox (1996), consider the range of potential causes leading to correct, illusory conjunction, and feature error responses. However, our point is simply this: Even without correcting for the difference in feature errors, there were more illusory conjunctions in the hard KF than in the hard RF condition. Given that there were, by experimental manipulation, more feature errors in the hard RF condition, any correction will increase the difference in binding errors even more.

These results suggest that increasing the difficulty associated with KF discrimination is more likely to produce binding errors than is similar difficulty in the RF. This conclusion is supported by comparing the hard KF and hard RF conditions in this experiment with the analogous shape/color condition of [Experiment 1](#). By comparing the condition in [Experiment 1](#) with the hard KF conditions, we can see the effect of increasing KF difficulty while holding RF difficulty constant. Here, we saw an increase in binding error rate from 11 % to 17.5 %, $F(1, 18) = 5.32$, $p = .033$. By comparing the condition in [Experiment 1](#) with the hard RF conditions, we see the effect of increasing RF difficulty while holding KF difficulty constant. This did not significantly change the binding error rate (11 % and 7 %), $F(1, 18) = 2.21$, $p = .15$. We conclude that there is an important asymmetry in the effects of feature discriminability on binding. Changing the KF, the attribute that defines the target,

changes the binding error rate. However, changing the RF has no apparent effect on binding efficiency.

General discussion

Our experiments have shown that what seems to matter in cross-dimension visual feature binding is not the feature dimensions themselves, but the ease with which the target-defining feature can be selected. In [Experiment 1](#), we controlled for the difficulty of KF and RF discriminations across different features, of color, orientation, and shape. To reiterate, the KF distinguished the target from the nontarget (e.g., color when red and green items were attended to and blue ones were ignored); the RF was the feature of the target to be reported (e.g., “S” when the shape of a red or green target was reported). When the difficulty of these discriminations was controlled for, the rate of binding errors was constant across the six different cross-dimensional binding tasks. However, in [Experiment 2](#), we manipulated the relative difficulty of the KF and RF discriminations. We then found that difficulty on the KF, but not the RF, increased binding error rates. Results from [Experiment 2](#) are restricted to the color/shape features. A similar approach can be used in future research to address the effects of difficulty associated to a wide range of features, including motion.

The color, shape, and orientation feature we chose had practical benefits, since they were easy to join in various combinations. But these features also have interesting differences and unique qualities. Color is a surface feature that can flexibly “fill in” contours, unlike shape or orientation (Treisman & Schmidt, 1982). Shape is a more complex form of feature than is orientation. Furthermore, a variety of findings and models consider color, shape, and orientation to be more or less coupled within the neural codes of the visual system. For example, the classic effect of McCollough (1965), showing orientation-specific color aftereffects, suggests some early binding of color and orientation, and significant numbers of neurons in the early visual cortex have receptive fields including orientation and color information (Hubel & Wiesel, 1968). Other models suggest a separation of *ventral* codes for attributes like color and other surface-based features and *dorsal* codes for attributes like orientation (Milner & Goodale, 1996). For example, damage to ventral areas resulting in achromatopsia can leave orientation and shape recognition relatively intact (Heywood, Wilson, & Cowey, 1987). Similarly, some accounts emphasize the way in which visual orientation may directly drive action, perhaps through the dorsal stream, in a different way from the associations of action and complex object shape, perhaps as processed by the ventral stream (Symes, Ellis, & Tucker, 2007).

Our results do not speak to different models of the association between feature dimensions, such as those above, within their specific context of application. The apparent

contradictions between them do not apply if we assume a variety of context-dependent feature codes. Color and orientation may be combined in the early visual cortex and yet be separate within the context of visual action and recognition. Letter shapes (such as those we used) may be primarily ventral, while orientation information may be spread more thoroughly across the visual system. So although many models might be used to generate diverse and even conflicting predictions about cross-dimensional feature binding, our results show something different. That is, any couplings of feature dimensions, available for other purposes within the visual system, are not available for our simple feature-binding task. Instead, it appears that for the purposes of feature binding, different feature dimensions are represented on a “separate but equal” basis. Our hypothesis is that efficiency of binding is unaffected by the particular pair of dimensions used. Here, there was no detectable bias for one feature dimension to bind more easily with another. The use of motion as stimulus would be an interesting test of this hypothesis.

These results are consistent with those of previous studies by Duncan (1993) showing that the costs of multiple discriminations did not depend upon the specific visual features being judged. The focus of Duncan’s study was quite different from ours. His study did not directly examine binding and did not address measurements of illusory conjunctions, feature errors, or binding error rates. However, his conclusion, that attention to an object provides access to all object features, is consistent with our findings.

Our conclusions are possible only because we maintained the difficulty of the different feature discriminations. By manipulating the difficulty of these discriminations, we gained additional general insights into cross-dimension feature binding. The effects of KF difficulty were perfectly consistent with models like FIT (Treisman, 1988) and the multiple-stage account (Humphreys et al., 2009). In both accounts, focused attention to a location or object plays a critical role in binding. Focused attention to the target is possible only if the target-defining feature can be selected. When it is difficult to select the target from the distractor, focused attention to the target is also difficult to achieve. In the absence of focused attention, it is well established that illusory conjunctions are more frequent (e.g., Treisman & Schmidt, 1982). In this way, we would naturally expect KF difficulty to affect binding error rates.

Although difficulty on the RF (by intention) increased feature errors, it did not increase binding error rates. This suggests that, once the target is selected on the basis of the KF, data-driven limits can make it difficult to identify the features of an object, without affecting how accurately features are assigned to objects. Put another way, making the feature report difficult might produce a noisy or degraded perception of the feature value, but this degraded representation is as easy to bind to the target object as a nondegraded one. This result is consistent with the FIT notion of a master

map (Treisman, 1988), in which feature values present in the selected location of the map are combined. The multistage model (Humphreys, 2001), by contrast, suggests a more interactive approach, in which bottom-up activations drive spatial attention to a location and top-down mechanisms act to stabilize early feature conjunctions. Our results are also consistent with this model, if it is assumed, as seems reasonable, that the bottom-up activations are not significantly reduced simply because discrimination is difficult to make. For example, when orientation was the RF and this was made to be a difficult discrimination, this meant that the target and distractor had similar orientation. However, both objects were still readily visible, and a reasonable interpretation according to the multistage model is that orientation had little effect on how efficiently the objects called attention to their location.

Finally, we would like to consider how one of our central results might relate to future neuropsychological and neurophysiological studies. We found that feature dimensions were “separate but equal” in terms of how easily they could be bound together; that is, there was no bias favoring some feature combinations over others, in the normal system. However, it is less clear what might happen in a lesioned brain. One possibility is a network of brain structures, which, overall, shows the “separate but equal” organization we observed but, at a more detailed level, reveals biases in feature combinations. For example, the pulvinar seems important for visual feature binding (Arend, Rafal, & Ward, 2008; Ward et al., 2002), and within the pulvinar, there is segregation of connections from all over the visual cortex. For example, connections with parietal and ventral areas are somewhat distinct. Selective lesions to such a structure could conceivably produce specific deficits in cross-dimension binding. Future research could build on the work here, by comparing, across a range of defining and probe dimensions, the binding errors following damage to brain areas involved in feature binding. The resulting picture would inform us about how maps in different parts of the brain may be specialized yet jointly contribute to coding and binding of different feature dimensions.

References

- Arend, I., Rafal, R., & Ward, R. (2008). Spatial and temporal deficits are regionally dissociable in patients with pulvinar lesions. *Brain*, 131, 2140–2152.
- Ashby, F. G., Prinzmetal, W., Ivry, R., & Maddox, T. (1996). A formal theory of illusory conjunctions. *Psychological Review*, 103, 165–192.
- Botella, J., Barriopedro, M., & Suero, M. (2001). A model of the formation of illusory conjunctions in the time domain. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 1452–1467.
- Cohen, A., & Ivry, R. (1989). Illusory conjunctions inside and outside the focus of attention. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 650–663.
- Cohen, A., & Rafal, R. D. (1991). Attention and feature integration: Illusory conjunctions in a patient with a parietal lobe lesion. *Psychological Science*, 2, 106–110.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293, 2470–2473.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 501–517.
- Duncan, J. (1993). Similarity between concurrent visual discriminations: Dimensions and objects. *Perception & Psychophysics*, 54, 425–430.
- Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: Recognition, navigation, or encoding? *Neuron*, 23, 115–125.
- Grabowecky, M., Robertson, L. C., & Treisman, A. (1993). Preattentive processes guide visual search: Evidence from patients with unilateral visual neglect. *Journal of Cognitive Neuroscience*, 5, 288–302.
- Heywood, C. A., Wilson, B., & Cowey, A. (1987). A case study of cortical color “blindness” with relatively intact achromatic discrimination. *Journal of Neurology, Neurosurgery & Psychiatry*, 50, 22–29.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195, 215–243.
- Humphreys, G. W. (2001). A multi-stage account of binding in vision: Neuropsychological evidence. *Visual Cognition*, 8, 381–410.
- Humphreys, G. W., Cinel, C., Wolfe, J., Olson, A., & Klempen, N. (2000). Fractionating the binding process: Neuropsychological evidence distinguishing binding of form from binding of surface features. *Vision Research*, 40, 1569–1596.
- Humphreys, G. W., Hodsoll, J., & Riddoch, M. J. (2009). Fractionating the binding process: Neuropsychological evidence from reversed search efficiencies. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 627–647.
- Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, 240, 740–749.
- McCollough, C. (1965). Color adaptation of edge-detectors in the human visual system. *Science*, 149, 1115–1116.
- Milner, A. D., & Goodale, M. A. (1996). *The visual brain in action*. Oxford: Oxford University Press.
- Milner, A. D., & Goodale, M. A. (2008). Two visual systems reviewed. *Neuropsychologia*, 46, 774–785.
- Mooney, C. Z. (1997). *Monte Carlo simulation*. Thousand Oaks, CA: Sage.
- Prinzmetal, W., Henderson, D., & Ivry, R. (1995). Loosening the constraints on illusory conjunctions: Assessing the roles of exposure duration and attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1362–1375.
- Prinzmetal, W., Ivry, R. B., Beck, D., & Shimizu, N. (2002). A measurement theory of illusory conjunctions. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 251–269.
- Quinlan, P. T. (2003). Visual feature integration theory: Past, present, and future. *Psychological Bulletin*, 129, 643–673.
- Sasaki, Y., & Watanabe, T. (2004). The primary visual cortex fills in color. *Proceedings of the National Academy of Sciences*, 101, 18251–18256.
- Symes, E., Ellis, R., & Tucker, M. (2007). Visual object affordances: Object orientation. *Acta Psychologica*, 124, 238–255.

- Tanaka, K., Saito, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology*, 66, 170–189.
- Treisman, A. M. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology*, 40A, 201–237.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Treisman, A. M., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 459–478.
- Treisman, A. M., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107–141.
- Ward, R., Danziger, S., Owen, V., & Rafal, R. (2002). Deficits in spatial coding and feature binding following damage to spatio-topic maps in the human pulvinar. *Nature Neuroscience*, 5, 99–100.
- Wei, P., Muller, H. J., Pollman, S., & Zhou, X. (2011). Neural correlates of binding features within- or cross-dimensions in visual conjunction search: An fMRI study. *NeuroImage*, 57, 235–241.
- Wolford, G., & Shum, K. H. (1980). Evidence for feature perturbations. *Perception & Psychophysics*, 27, 409–420.