What Do We Learn From Binding Features?
Evidence for Multilevel Feature Integration

Lorenza S. Colzato
Leiden University

Antonino Raffone
University of Sunderland and
RIKEN Brain Science Institute

Bernhard Hommel
Leiden University

Four experiments were conducted to investigate the relationship between the binding of visual features (as measured by their aftereffects on subsequent binding) and the learning of feature–conjunction probabilities. Both binding and learning effects were obtained, but they did not interact. Interestingly, (shape–color) binding effects disappeared with increasing practice, presumably because of the fact that only 1 of the features involved was relevant to the task. However, this instability was only observed for arbitrary, not highly overlearned combinations of simple geometric features and not for real objects (colored pictures of a banana and strawberry), where binding effects were strong and resistant to practice. These findings suggest that learning has no direct impact on the strength or resistance of bindings or on speed with which features are bound; however, learning does affect the amount of attention particular feature dimensions attract, which again can influence which features are considered in binding.

Keywords: feature integration, learning, binding problem, attentional set, event file

Considerable evidence suggests that cortical networks encode the external environment in a distributed fashion. A striking example of spatially distributed coding in cortical information processing is given by the primate visual cortex, processing visual event features in parallel in numerous cortical maps (Cowey, 1985; Felleman & van Essen, 1991). This coding scheme also applies to events in the auditory and other sensory modalities and to multimodal event processing. As handy as distributed coding may be, it creates numerous so-called "binding problems," that is, difficulties in relating the codes of a given entity or processing unit (e.g., visual object) to each other. To resolve these problems, the brain needs some sort of integration mechanism that binds together the distributed codes belonging to the same event, while keeping these codes separated from codes for other events (Treisman, 1996).

Mechanisms of Feature Integration

At a neural level, a theoretical solution to the binding problem may be given by high-order cardinal cells (Barlow, 1972), onto which signals from neurons coding for the to-be-bound features converge. However, given the high variability of objects belonging to a given category in terms of their instances and retinal projections, as well as the numerous ways in which discrete features can be potentially combined, the exclusive reliance on convergent mechanisms would ultimately lead to a combinatorial explosion and is therefore not plausible. Another potential solution to the binding problem is given by cell (neural) assemblies or sets of tightly connected neurons, the identity of the assembly being defined in terms of higher firing rates or coactivation of the participating neurons (Amit, 1995; Braitenberg & Schüz, 1991; Hebb, 1949). In this representational scheme, individual neurons encode for simple features, and the associative connections between these neurons enable pattern encoding and completion within the assembly. This solution avoids the combinatorial explosion problem implied by cardinal cells and, thus, seems to be well suited for arbitrary, frequently changing feature combinations.

At a behavioral level, one way to study feature binding mechanisms is to put processing systems under conditions that render proper integration difficult or impossible and then to look for the creation of incorrect bindings or "illusory conjunctions" (Treisman & Schmidt, 1982). Another way is to search for aftereffects of feature integration, that is, for side effects of created feature bindings on later performance. In a seminal study along these lines, Kahneman, Treisman, and Gibbs (1992) presented participants with two displays in a sequence, a brief multiletter prime

display (S1) followed by a single-letter probe display (S2) requiring verbal identification. Having just seen the probe letter somewhere in the prime display, the subjects tended to facilitate probe identification. However, more reliable than this nonspecific repetition effect was the benefit of repeating the particular combination of stimulus and location, a finding that since then has been replicated many times in both the visual (e.g., Gordon & Irwin, 1996; Henderson, 1994; Henderson & Anes, 1994; Hommel, 1998; Park & Kanwisher, 1994) and the auditory modality (Mondor, Hurlburt & Thorne, 2003). Interestingly, binding aftereffects are not restricted to stimulus shape and location but also can be found for other feature combinations, such as shape and color (Hommel, 1998; Hommel & Colzato, 2004)—the features on which the present study will focus.

Apparently, perceiving an event automatically creates a kind of "object file" (Kahneman et al., 1992) or "event file" (Hommel, 1998), an integrated episodic trace containing information about the various features and bindings of that event. Reviewing parts or aspects of that event automatically retrieves the file, which produces a benefit if previous and present event perfectly match (Kahneman et al., 1992) and/or code confusion if they mismatch (Hommel, 1998, 2004). Indeed, incomplete repetitions (e.g., shape match combined with color mismatch) commonly produce worse performance than conditions in which no stimulus feature is repeated, whereas the latter yield performance comparable with complete repetitions (Hommel, 1998; Hommel & Colzato, 2004). That is, reusing an already created object file seems to be of little help but retrieving an old file that also includes mismatching codes apparently causes conflict (Hommel, 2004).

Taken altogether, the available evidence strongly suggests that perceiving an event results in the integration of its features, that is, in the binding of the individual codes representing them. Once bound, the feature codes can no longer be selectively addressed, so that perceiving some combination of the same features retrieves the whole file, a kind of pattern-completion process. Feature binding is supposed to be a fast-acting process (simple bindings emerge after 300 ms or less: Hommel & Colzato, 2004) that creates transient representational structures. In the present study we asked how this process might be related to the learning of feature combinations—that is, the creation of relatively permanent memory changes.

Binding and Learning

On the one hand, one may consider conjunction learning being a direct consequence of binding—we call this the *strong-dependence hypothesis*. As suggested by Fell, Fernandez, Klaver, Elger, and Fries (2003), synchronized neural activity may cause Hebbian learning (neurons that fire together, wire together), that is, learning through the long-term modification of synaptic efficacy induced by reverberation of neural activity in cerebral circuits (Hebb, 1949). Indeed, Miltner, Braun, Arnold, White, and Taub (1999) demonstrated that associative learning in humans is accompanied by neural synchronization between the brain areas representing the to-be-associated stimuli. Along these lines, one would expect that binding particular features is a first, preliminary step toward creating a more durable memory trace, suggesting that relevance and impact of binding decrease over time, to the degree that feature conjunctions approach their asymptotic association values. If

binding and conjunction learning go hand in hand, learning a given feature conjunction should affect the way in which these features are bound, in terms of *strength* and *speed* of binding.

On the other hand, however, binding and learning are expressed over different time-scales and mediate different kinds of neural representations, that is, perceptual and active working memory representations versus latent long-term memory representations. Binding processes are thought to solve problems resulting from distributed processing, whereas learning processes concern the long-term storage of information that is likely to be retrieved and used on a later occasion. Although some combinations of features are more likely than others, many feature conjunctions are so arbitrary—just think of the color of a shirt or the font of a letter—that it makes little sense to perpetuate them by creating a long-lasting memory. That is, not much of what binding processes integrate is worthwhile to maintain for much longer than the event in question is perceptually available, which leaves the possibility that binding and learning are less tightly connected than the strong-dependence hypothesis suggests. If so, one would not expect a significant impact of long-term learning on the effects of short-term binding with arbitrary feature combinations. Conversely, on this *weak-dependence hypothesis* one would expect that the learning rate for conjunctive code transfer from short-term binding to long-term learning is relatively low, thus enabling one to filter-out conjunction occurrences with a low behavioral salience.

A strong-dependence hypothesis would suggest that long-term learning-related factors such as object familiarity, repetition (frequency) of stimulus feature-conjunctions, and the frequency of association between stimulus feature-conjunctions and responses in a given task setting modulate short-term binding effects in terms of response times or response accuracy. By contrast, a weak-dependence hypothesis would predict a negligible modulation of conjunction learning or associations of perceptual binding codes to responses upon binding-related effects. To test these alternative hypotheses, we conducted four experiments in which participants were presented with two stimuli in a row, S1 and S2. These stimuli varied on two dimensions, shape and color, thus creating a set of four possible feature combinations. To avoid confounding stimulus repetition effects with response repetition effects we used Hommel's (1998) experimental design, which comprises a precued left-right response (R1) to the mere onset of S1 (so that no S1 feature was correlated with a particular response and S1 repetitions and R1 repetitions were independent) and a left-right response (R2) to the shape of S2. The general idea was to make two of the four feature combinations more likely than the other two to induce strong associations between the underlying codes. If binding effects strongly depended upon previous learning, a greater number of presentations of a particular feature conjunction (conjunction learning strength) or familiarity of a given conjunction (object) should affect the way the respective features are bound, by either reducing or boosting the impact of this binding on different aspects of performance.

Experiment 1

Experiment 1 was modeled after Hommel (1998): Participants were cued to prepare a left- or right-hand keypress (R1), which they carried out as soon as S1—the *prime stimulus*—was presented (Figure 1). Although the identity of S1 did not matter for the
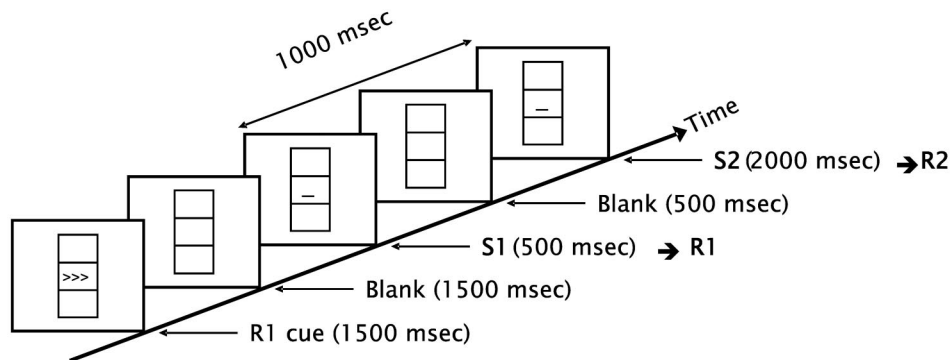
Figure 1. Overview of the displays and the timing of events in Experiments 1 and 2.

response, it varied in shape or orientation (horizontal vs. vertical line) and color (red vs. green). One second later, S2 appeared to determine R2. The two alternative shapes of S2 were mapped onto the two R2 alternatives, whereas the color of S2 was entirely irrelevant to the task. Our focus was on interactions between shape (orientation) and color-repetition effects. On the basis of earlier findings (Hommel, 1998; Hommel & Colzato, 2004), we expected that shape repetitions produce better performance on S2 than shape alternations if color is also repeated, but worse performance if color alternates. In other words, performance on S2 should be better with a complete S1-S2 match or mismatch than for partial matches—a pattern that we will call *partial-repetition cost*. We hypothesize that shape and color features of S1 are still bound when processing S2, so that repeating one feature of S1 would also reactivate the other one, causing an increased interference at S2-dependent response selection.

The crucial question was whether this interaction would vary as a function of conjunction learning or, more precisely, as a function of the relative frequency (i.e., probability) of a given feature conjunction. In Experiment 1 we manipulated the conjunction frequency by presenting two shape-color combinations of S1 (e.g., green-vertical and red-horizontal) four times as often as the other two (red-vertical and green-horizontal). This manipulation was assumed to induce stronger associations between the codes of the more frequent pairs of features, which should yield a main effect of frequency. But more important than this main effect was whether frequency would interact with our measure of binding (as the strong- but not the weak-dependency hypothesis would predict), that is, whether frequency would modulate the interaction between shape repetition and color repetition (the partial repetition cost).

This frequency manipulation also may affect other targets than the targeted aspects of performance. In particular, introducing unbalanced frequencies will raise particular expectations leading to a higher degree of preparedness or bias of the cognitive system toward the more probable stimuli. To separate these context-bound short-term effects from the impact of a proper long-term learning, we ran two blocks: an acquisition block in which the conjunction-frequency manipulation was administered and a test block in which all feature combinations were equally probable. The critical test was whether running through the acquisition block would affect performance (i.e., would modulate the interaction between shape repetition and color repetition) in the test block. However,

working through an extended block of trials may also have effects that are unrelated to the frequency manipulation proper. To control for such nonspecific effects we compared performance in the experimental group (where frequencies were unbalanced in the acquisition block as described) with that of a control group (where frequencies were balanced).

Methods

Participants. Thirty-six students of the Leiden University took part for pay in Experiment 1, 18 in the experimental and 18 in the control group. All reported having normal or corrected-to-normal vision and were not familiar with the purpose of the experiment.

Apparatus and stimuli. The experiments were controlled by a Targa Pentium III computer, attached to a Targa TM 1769-A 17″ monitor. Participants faced three gray square outlines, vertically arranged, as illustrated in Figure 1 (the top and bottom frames served no purpose in the present study but were kept for comparison with other experiments from our laboratory). From viewing distance of about 60 cm, each of these frames measured 2.6° × 3.1°. A vertical line (0.1° × 0.6°) and a horizontal line (0.3° × 0.1°) served as S1 and S2 alternatives, which were presented in red or green in the middle frame. Response cues were also presented in the middle frame (see Figure 1), with rows of three left- or right-pointing arrows indicating a left and right keypress, respectively. Responses to S1 and to S2 were made by pressing the left or right shift key of the computer-keyboard with the corresponding index finger.

Procedure and design. The experiment consisted of a 1-hr *acquisition* session and a 30-min *test* session. In both sessions participants carried out two responses per trial. R1 was a simple reaction with the left or right key, as indicated by the response cue. It had to be completed as soon as S1 appeared, independent of its shape or color. Participants were informed that there would be no systematic relationship between S1 and R1, or between S1 and S2, and they were encouraged to respond to the *onset* of S1 only, disregarding the stimulus' attributes. R2 was a binary-choice reaction to S2. Half of the participants responded to the vertical and the horizontal line by pressing the left and right key, respectively, whereas the other half received the opposite mapping. All the participants began with the acquisition session and after a 5-min break continued with the *test* session. Half of the participants (*control group*) received a *balanced* acquisition session (in which every features combination of S1 had the same probability to occur) whereas the other half (*experimental group*) received an *unbalanced* acquisition session (in which we manipulated the familiarity to occur of S1: half of these participants received as S1 for 40% of the time the combination of horizontal and red line and of vertical and green line and for 10% of the time the combination of horizontal and green line and of vertical and

red line, whereas the other half received the opposite mapping).[1] The test session was the same for both groups: every features combination of S1 had the same probability to occur.

The sequence of events in each trial is shown in Figure 1. A response cue signaled a left or right key press (R1) that was to be delayed until presentation of S1, a red or green, vertical or horizontal line in the middle box. S2 appeared 1 s later at the same location—another red or green, vertical or horizontal line. S2 shape signaled R2, also a speeded left or right key press. R2 speed and accuracy were analyzed as function of the repetition versus alternation of stimulus shape and color. If the response was incorrect auditory feedback was presented.

The acquisition phase comprised 320 trials composed by a factorial combination of the two shapes (vertical vs. horizontal line), colors (red vs. green) of S2, the repetition versus alternation of shape and color and, only for the unbalanced phase, the frequency (high = 80% vs. low = 20%) of S1. In the balanced acquisition phase, every feature combination was repeated 20 times whereas in the unbalanced acquisition session the high-frequency trials were repeated 32 times and the low-frequency trials only 8 times. The test session comprised 224 trials composed by the same factorial combination as in the acquisition session except for the frequency manipulation: every features combination had the same probability to appear and was repeated 14 times. Thus, taken together, the two sessions amounted to 544 trials.

Results

S1-R1. The significance criterion for all analyses was set to $p < .05$. We first analyzed the data from R1, the prepared response to S1. In case of errors or anticipatory responses (reaction times [RTs] <500 ms) subjects had to repeat R1 immediately. Mean correct RTs were analyzed as a function of conjunction frequency (high vs. low, dummy-coded for the control group) and session (test vs. acquisition), the two within-participant factors, and group (control vs. experimental) as between-participants factor. There was only a main effect of session, $F(1, 34) = 30.10, p < .001$, indicating that responses sped up from the acquisition phase (347 ms) to the test session (299 ms).

S2-R2. After excluding trials with missing or anticipatory responses (1.6%), mean RTs and proportions of errors (PEs) for R2 (i.e., the response to S2) were analyzed as a function of group, frequency, session, and the possible relationships between S1 and S2, that is, repetition versus alternation of stimulus shape or color (see Table 1 for means).

The RTs produced three reliable effects: a main effect of session, $F(1, 34) = 4.14, p < .05$, indicating faster responses in the test session, and a two-way interaction of color and shape, $F(1, 34) = 6.30, p < .05$, that was modified by a three-way interaction with session, $F(1, 34) = 9.85, p < .005$. Separate analysis of variances (ANOVAs) for the two sessions revealed that shape and color interacted in the acquisition session, $F(1, 34) = 16.64, p < .001$, but not in the test session, $p > .7$. The acquisition session exhibited typical partial-repetition costs, that is, better performance for shape repetition than alternation if color is repeated but worse performance for shape repetition than alternation if color alternates (for the experimental group, see Figure 2). That is, inasmuch as shape–color interactions are produced by feature binding, these features are less likely to get bound as practice increases. Importantly, this effect does not seem to be caused by or related to the frequency manipulation, as indicated by the absence of higher-order interactions involving group, $p > .8$ (shape × color × frequency × session), $p > .7$ (shape × color × fre-

quency × group), and $p > .5$ (shape × color × frequency × group × session).

PEs mirrored the RTs and produced two reliable effects: a main effect of session, $F(1, 34) = 5.81, p < .05$, and a three-way interaction that included session, shape, and color, $F(1, 34) = 4.66, p < .05$. Subjects made fewer errors (7.5%) in the test than in the acquisition session (9.4%). As for the RTs, separate ANOVA for the acquisition and the test session revealed that the former exhibited typical partial-repetition costs, which were gone in the latter.

Discussion

Experiment 1 produced three important results. First, as expected on the basis of earlier observations (Hommel, 1998; Hommel & Colzato, 2004), there was evidence that shape and color were integrated more or less automatically. Thus, although color was a response-irrelevant dimension in the task, color codes engaged in shape-color binding.

Second, this shape-color binding effect disappeared as practice increased, suggesting that it is unstable for some reason. One possibility is that practice is accompanied by a fine-tuning of selective attention to stimulus dimensions. Because color is not relevant to the task, stimulus color may attract some attention in the beginning of the task but lose impact over the course of time. Indeed, manipulations of task relevance have provided evidence that feature integration is modulated by attention to feature dimensions (Hommel, 1998), suggesting that the attentional set determines *what* gets integrated rather than *whether* integration takes place (Hommel, 2004).

Third, there was no evidence that feature combinations that are more probable increase or decrease (i.e., modulate) the aftereffects of color-shape integration. This may be taken to provide evidence that conjunction learning and short-term feature binding are mediated by different mechanisms. However, there are two reasons to be careful with respect to such a conclusion. One reason is obvious from Figure 2: For infrequent combinations, the interaction between color and shape in the acquisition phase disappears entirely in the test phase. For frequent combinations, however, there is

---

[1] In this design, the manipulations of feature-combination frequency on the one hand and feature overlap on the other are entirely orthogonal; for example, if the combinations of horizontal-red (*Hr*) and vertical-green (*Vg*) are frequent and combinations of horizontal-green (Hg) and vertical-red (Vr) infrequent, the (S1-S2) sequences of Vr-*Hr*, Hg-*Hr*, Vr-*Vg*, and Hg-*Vg* would be the partial repetitions of the frequent S2 (leaving *Hr-Hr* and *Vg-Vg* as complete repetitions and *Vg-Hr* and *Hr-Vg* as alternations), whereas *Hr*-Vr, *Vg*-Vr, *Hr*-Hg, and *Vg*-Hg would be the partial repetitions of the infrequent S2 (leaving Vr-Vr and Hg-Hg as complete repetitions and Hg-Vr and Vr-Hg as alternations). We note that this introduces a contingency with respect to the *sequence* of frequent and infrequent combinations: If S2 is frequent, all complete repetitions or alternations are sequences of the type "frequent–frequent" and all partial repetitions are sequences of the type "infrequent–frequent"; if S2 is infrequent, all complete repetitions or alternations are sequences of the type "infrequent-infrequent" and all partial repetitions are sequences of the type "frequent–infrequent." Given the results of the present study this contingency is unlikely to be relevant in this context, but one may consider future outcome patterns for which it could provide an interesting point of departure for explanation. We are grateful to Bruce Milliken for pointing this out.

Table 1

Acquisition and Test Sessions in Experiments 1 and 2: Means of Mean Reaction Times (RTs) for Responses to Stimulus 2 (in ms) and Percentages of Errors (PEs) on R2, as a Function of Group, Frequency of Feature Conjunctions, and the Feature Match Between Stimulus 1 and Stimulus 2

	Experiment 1								Experiment 2			
	Control group				Experimental group				Experimental group			
	Frequent		Infrequent		Frequent		Infrequent		Frequent		Infrequent	
Match	RT	PE	RT	PE	RT	PE	RT	PE	RT	PE	RT	PE
Acquisition session												
Neither	521	8.3	526	9.7	497	6.2	500	5.2	463	5.7	492	8.7
S(hape)	542	10.2	530	8.3	505	6.3	498	4.9	470	6.6	510	10.8
C(olor)	540	10.0	537	9.4	511	6.3	519	5.2	469	7.5	493	9.0
SC	526	8.7	523	11.7	499	5.8	496	3.5	460	6.6	472	8.6
Test session												
Neither	496	14.2	511	9.3	482	6.5	489	7.9	474	5.1	469	6.9
S(hape)	508	12.5	503	9.1	484	5.7	489	5.9	476	7.1	461	5.9
C(olor)	491	10.9	505	10.5	494	7.7	483	7.7	473	7.7	467	8.1
SC	502	14.9	508	13.1	477	6.7	482	7.7	449	6.1	463	5.3

hardly any change in the interaction from acquisition to test—even if it is statistically reliable in the former but not the latter. Hence, the qualitative pattern might suggest that more learning makes the effect of binding more robust over time. The statistics point in the same direction: In a separate ANOVA on the high frequency conditions the four-way interaction of color and shape with session and group reached the 13% level.

Figure 2. Means of mean reaction times and percentage of errors for responses to the second stimulus (S2→R2) in Experiment 1 (experimental group only), as a function of the match between (i.e., repetition vs. alternation of) S1 and S2 with respect to shape and color, of the frequency of the shape-color conjunction (in S2), and session (acquisition vs. test).

Another reason for not jumping to conclusions is that we failed to find any systematic impact of the frequency manipulation. Given that participants did not need to identify S1 (which shows in the rather fast RTs) it may not be too surprising that frequency did not affect R1 performance. However, we had hoped that the frequency manipulation on S1 would transfer to S2 and, hence, R2, given that S1 and S2 were made up of the same features. Apparently, this transfer did not take place, perhaps because conjunction learning is too context-sensitive. Whatever the reason for the lack of a frequency effect, it renders the absence of an interaction between binding and frequency-induced learning uninformative. To accommodate that, in Experiment 2 we reran the experimental group but this time manipulated the frequency of feature conjunctions in S2. As we will see, this manipulation was successful in creating a frequency effect on R2. After this demonstration, Experiment 3 will address the remaining surprise—the lack of transfer from the acquisition to the test phase.

Experiment 2

As pointed out, one reason for why we failed to find an interaction between learning and aftereffects of binding in Experiment 1 might have to do with the fact that we manipulated the frequency of S1 but not S2: stimulus frequency may not (strongly) affect simple responses or its impact may not transfer from S1-R1 to S2-R2. Accordingly, Experiment 2 was conducted, in which the frequency manipulation concerned S2 instead of S1. This should produce a main effect of frequency on R2 and, if the lack of transfer was indeed the critical factor, the sought-for interaction of binding aftereffects and learning in terms of biased conjunction frequencies.

Methods

Sixteen new students participated, which fulfilled the same criteria as in Experiment 1. The method was exactly as in the experimental group of Experiment 1, except that the frequency manipulation referred to S2
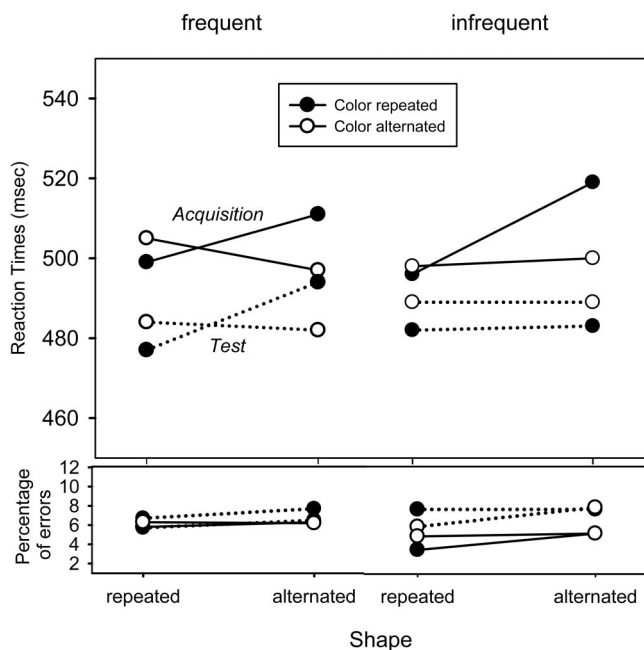
instead of S1. The control group from Experiment 1 was used for comparison.

Results

The data were analyzed analogously to Experiment 1. The control group from Experiment 1 was included to create the Group factor.

S1-R1. Mean correct RTs were analyzed as a function of session (test vs. acquisition) and group (control [from Experiment 1] vs. experimental). The only significant effect of session, $F(1, 34) = 46.92$, $p < .001$, indicated that responses were faster in the test phase (272 ms) than in the acquisition phase (312 ms).

S2-R2. Trials with missing or anticipatory responses (1.8%) were excluded from the analysis. Table 1 provides an overview of the means for RTs and PEs obtained for R2.

The RT analysis produced two clusters of effects. One cluster involved a main effect of session, $F(1, 34) = 14.68$, $p < .001$, and frequency, $F(1, 34) = 7.13$, $p < .05$, and a three-way interaction including session, frequency, and group, $F(1, 34) = 14.43$, $p < .001$. Apart from a general practice effect, the underlying pattern revealed that the frequency effect was restricted to the acquisition phase of the experimental group (466 vs. 492 ms, i.e., a 26-ms benefit for frequent feature combinations) but absent in the test phase (468 vs. 465 ms) and in both phases of the control group (532 vs. 529 ms and 499 vs. 507 ms). That is, the frequency manipulation was successful in affecting R2 performance, even though it failed to transfer to the test phase.

The other cluster comprised a color main effect, $F(1, 34) = 5.07$, $p < .05$, the expected interaction of shape and color, $F(1, 34) = 17.67$, $p < .001$, which was further modified by session, $F(1, 34) = 8.55$, $p < .01$. As shown in Figure 3 (for the experimental group), the acquisition phase produced the common partial–repetition–cost pattern whereas the test phase did not—an observation that was confirmed by separate ANOVAs, $F(1, 34) = 28.96$, $p < .001$ and p .5, respectively. Thus, as in Experiment 1, the shape–color interaction and the feature binding process it indicates seem to disappear with practice. However, as in Experiment 1, there was again an (again unreliable) indication that practice might affect binding in the high-frequency condition to a lesser degree than under high frequency (see Figure 3, dotted lines): whereas the two four-way interactions involving shape, color, and frequency were far from significant, $p > .30$ (shape × color × frequency × session), $p > .48$ (shape × color × frequency × group), the five-way interaction reached the 16% level.

The errors yielded three two-way interactions: the interaction of session and group, $F(1, 34) = 5.92$, $p < .05$, indicated that the experimental group made more errors in the acquisition (7.9%) than in the test session (6.6%), whereas the control group showed the opposite tendency (9.6% vs. 11.8%). The interaction of frequency and group, $F(1, 34) = 7.25$, $p < .05$, reflects that frequency affected performance in the experimental group (6.6% vs. 7.9% for high and low frequent combinations, respectively) but not in the control group (11.2% vs. 10.1%). The session-by-frequency interaction, $F(1, 34) = 10.07$, $p < .005$, pointed to a somewhat larger and more positive impact of frequency in the acquisition session (8.0% vs. 9.5) than in the test session (9.8% vs. 8.6%). Finally, a three-way interaction of color, shape, and group, $F(1, 34) = 9.03$, $p < .005$, indicated that the interaction of color and
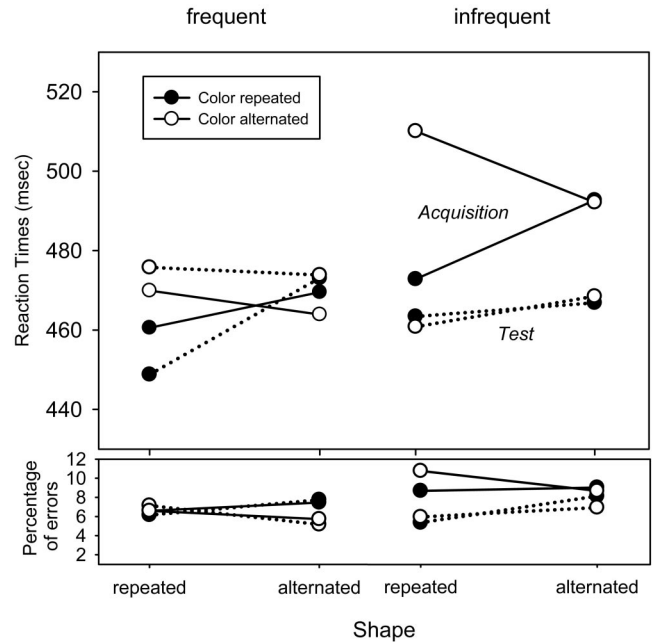

Figure 3. Means of mean reaction times and percentage of errors for responses to the second stimulus (S2→R2), in Experiment 2 (experimental group only), as a function of the match between (i.e., repetition vs. alternation of) S1 and S2 with respect to shape and color, of the frequency of the shape-color conjunction (in S2), and session (acquisition vs. test).

shape was reliable in the experimental group but not in the control group.

Discussion

Experiment 2 replicated the first two findings of Experiment 1, namely, shape and color repetitions interacted (i.e., produced partial-repetition costs) and this interaction disappeared with practice—reinforcing the idea of a progressive attentional tuning to the relevant stimulus dimensions. Even more important for present purposes, we also found evidence that the unbalanced frequency of feature combinations leads to an immediate adjustment that benefits performance, which confirms that our manipulation has worked. However, this benefit does not transfer to later performance on the same task (the test phase) and it does not seem to affect feature integration aftereffects more strongly than in Experiment 1. Thus, it seems unlikely that the lack of interaction between frequency manipulation and binding aftereffects in Experiment 1 was merely due to a lack of transfer from S1-R1 to S2-R2. And yet, both Experiments 1 and 2 revealed some signs that frequency might affect binding. In particular, it seems that more frequent and thus better learned conjunctions prevent shape-color binding from disappearing with practice. It is possible that this impact of learning was too weak in our experiments because even in the high frequency condition the number of repetitions was too low. Experiment 3 was intended to fix that possible problem by using highly overlearned feature conjunctions.

Experiment 3

On the basis of Experiments 1 and 2, the lack of interaction between conjunction learning and binding aftereffects may be the

result of two reasons. On the one hand, these two processes may work independently of each other, so that any further attempt to find the sought-for interactions would be doomed to fail. On the other hand, however, our manipulation of conjunction frequency may not have led to learning of a proper integrated representation at high levels of object representation (in inferotemporal cortex) but, rather, to a merely transient general bias toward the more likely conjunctions—by a kind of continuously updated situational model held in working memory (see Duncan, 2001). To address this possibility, we designed Experiment 3 to make sure that stably learned feature combinations in object representation are involved. Instead of the arbitrary feature conjunctions used in Experiments 1 and 2, Experiment 3 used images of real-life stimuli, a banana and a strawberry, which could appear in either their "natural" colors— that is, yellow and red—or in the opposite colors—red and yellow—which participants were unlikely to have experienced frequently in combination with these objects. That is, we used pre-experimental stimuli and presented them in either their standard color or in a color that is unlikely to be associated with them. Accordingly, we did not include an acquisition phase but had participants work through a test session only, where every feature combination of S1 and of S2 was equally probable. Analogously to the preceding experiments we hypothesized that familiar, that is, stably learned feature combinations, such as the yellow banana and the red strawberry, might affect binding differently than less familiar combinations, such as a red banana and a yellow strawberry.

Methods

The participants were 24 students who fulfilled the same criteria as those in Experiment 1. The experiment consisted of only one half-hour session. The procedure and the sequence of events were as in the test session of Experiments 1 and 2, with the following exceptions: Instead of a vertical and a horizontal line we presented figures of a banana ($0.3° \times 0.6°$) and a strawberry ($0.5° \times 0.6°$, see Figure 4), appearing in red or yellow inside the middle frame. Half of the participants responded to the shape of the banana and the strawberry by pressing the left and right key, respectively, while the other half received the opposite mapping. As in the previous experiments, color varied orthogonally to the shape and was completely irrelevant to the task.

Results

S1-R1. Correct RTs were analyzed as a function of the familiarity of the feature combinations. Familiarity indeed affected performance by producing faster responses with familiar than unfamiliar combinations (304 vs. 313 ms), $F(1, 22) = 4.46, p < .05$.

S2-R2. Data were analyzed as a function of familiarity, shape repetition, and color repetition. Trials with missing or anticipatory responses (1.3%) were excluded from the analysis. No reliable effect was obtained for PEs. In RTs, familiarity produced a main

Figure 4. Bitmaps of the stimuli used in Experiments 3 and 4.

effect, $F(1, 22) = 7.31, p < .05$, indicating that responses were faster to familiar than to unfamiliar combinations (498 vs. 511 ms). The only other significant effect was the interaction between shape and color repetition, $F(1, 22) = 9.10, p < .01$, following the standard pattern indicative of partial-repetition cost (see Figure 5). Importantly, however, there was not any hint to a three-way interaction involving familiarity, $F < 1$.

Discussion

Experiment 3 was successful in demonstrating a frequency effect induced by lifelong learning that affected performance even under conditions in which all feature combinations were equally probable—suggesting that in this experiment a different, probably "higher" level of object representation and perceptual learning was involved. This finding supports our speculation from the first two experiments that a sufficient degree of learning works against the disappearance of binding-related effects with practice. As another important result, we replicated the partial–repetition–cost pattern with real-life, object-like stimuli, which demonstrates that the previous observations with simple geometric elements (e.g., Hommel, 1998) are generalizable. In fact, the shape–color interaction was even more pronounced than observed with simpler stimuli, an issue we will address in Experiment 4. And, yet, there is still no evidence of any interdependency of learning and binding aftereffects. Even though our considerations are based on a null effect, which necessarily renders them preliminary, we thus tend to conclude that short-term binding aftereffects, like the partial repetition cost, and long-term learning, are independent processes.

Experiment 4

Introducing familiar objects in Experiment 3 produced both frequency effects and the common shape-color interaction without demonstrating the sought-for modulation of the latter through the former. In this respect, using familiar objects did not change the outcome obtained in Experiments 1 and 2 or the conclusion they suggested. However, we have mentioned that something did change in Experiment 3: the shape-color interaction was more pronounced. One possible (though theoretically less interesting) reason for this observation may relate to the amount of practice. As briefly considered in Experiment 1, practice on the task may allow fine-tuning input selection processes and thus increasingly prevent the irrelevant color information from being processed. In Experiments 1 and 2, fine-tuning could begin in the acquisition phase already, so that color would no longer be processed in the test phase. Given that there was no acquisition phase in Experiment 3 and the fact that the test block was somewhat shorter than the acquisition block, the stronger effect in Experiment 3 may thus simply reflect the fact that participants in Experiments 1 and 2 were tested after more extended practice.

However, the outcome of some post hoc analyses renders this account implausible. First we divided the acquisition trials in Experiments 1 and 2 (i.e., the two experimental groups and the control group) in four equal miniblocks and reran ANOVAs on these data with miniblock as additional factor. It turned out that the shape-color interaction interacted with miniblock, $F(3, 153) = 2.70, p < .05$. Separate analyses showed that shape and color interacted in the first miniblock only, $p < .001$, but not in the other
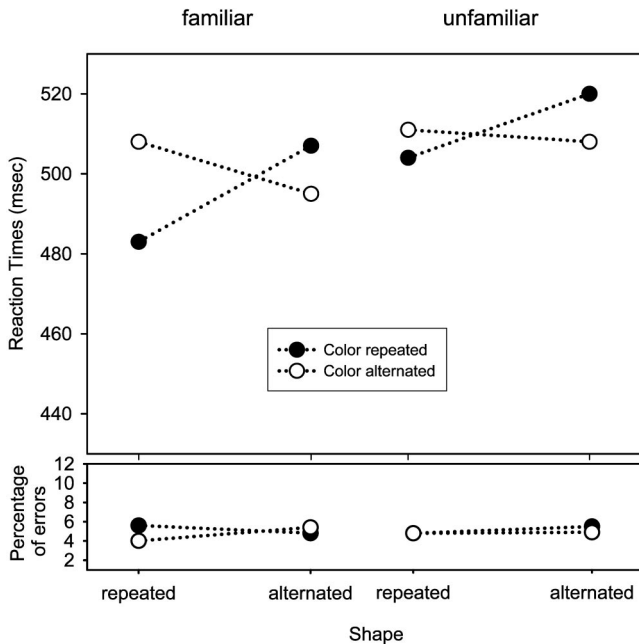
Figure 5. Means of mean reaction times and percentage of errors for responses to the second stimulus (S2→R2), in Experiment 1, as a function of the match between (i.e., repetition vs. alternation of) S1 and S2 with respect to shape and color, and of the familiarity (familiar vs. unfamiliar) of object-color conjunction (in S2).

miniblocks, $p > .62$, $p > .14$, and $p > .37$. The same analysis of the test trials in Experiment 3 (block length equated) did not reveal any modification of the shape-color interaction by miniblock, $F < 1$. That is, in Experiments 1 and 2, 80 trials of practice were sufficient to eliminate the shape-color interaction, which in Experiment 3 survived 224 trials without any drop in size. This conclusion was further confirmed by a direct comparison of performance in the control group of Experiment 1 and Experiment 3, in which we included the first two acquisition miniblocks from the former and the first two test miniblocks from the latter. The shape-color interaction was modified by miniblock and experiment, $F(1, 38) = 3.65$, $p < .05$. Separate analyses showed that experiment modified the shape-color interaction in the second, $p < .05$, but not the first miniblock, $p > .29$. That is, practice seems to eliminate shape-color integration—presumably by gating out color information—but only if the stimulus material consists of arbitrary geometric symbols.

Experiment 4 was designed to disentangle two possible interpretations of this outcome. Clearly, the stimulus material differed between Experiments 1 and 2 on the one hand and Experiment 3 on the other: the stimuli were more simple and arbitrary, and less "biological" in the former than the latter. This implies that the stimuli used in Experiment 3 may be cognitively and neurally represented in a different way than the lines used in the previous experiments. They are objects and are likely to be perceived and categorized as such, which among other things will involve the activation of conceptual traces in long-term memory—which, after all, was the reason to employ them. This fact may change the way these stimuli were processed, for reasons that we will elaborate in

the General Discussion section. However, before discussing this issue we need to consider another possible factor. Not only were the stimuli in Experiment 3 more object-like in terms of their more complex shapes and meanings, they also appeared in their standard colors—at least in 50% of the trials. To rule out that this color appearance was the responsible factor, we replicated Experiment 3 but replaced the two "biologically plausible" colors by colors that were unlikely to be closely associated with one or the other object.

Methods

Twenty-four students participated, they all fulfilled the same criteria as in Experiment 1. The procedure and the sequence of events were as in Experiment 3, except that the colors were pink and blue instead of yellow and red.

Results

S1-R1. Valid responses were performed in 280 ms on average.
S2-R2. The data were analyzed as in Experiment 3, except that the familiarity factor no longer applied. The errors yielded a main effect of shape repetition, $F(1, 22) = 10.92$, $p < .01$, as the result of more errors being made with shape repetitions than alternations (7.6% vs. 5.4%). However, as the shape effect had the opposite sign in RTs (459 vs. 465 ms, an unreliable difference) this might reflect a mere speed–accuracy trade-off. The RTs produced a main effect of color repetition, $F(1, 22) = 4.96$, $p < .05$, and an interaction of shape and color repetition, $F(1, 22) = 13.66$, $p < .01$. Whereas the former indicated a 7-ms benefit for color repetitions, the latter followed the expected cross-over cost pattern shown in Figure 6. An additional analysis with miniblock as a
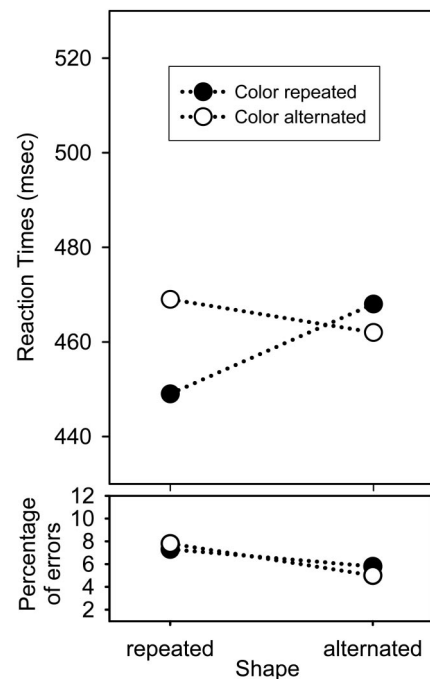

Figure 6. Reaction times and percentage of errors in Experiment 4, as a function of the repetition versus alternation of stimulus shape and stimulus color (in S2).

factor did not provide any evidence that the shape-color interaction decreases with practice, $F < 1$.

Discussion

In summary, we find the same outcome as in Experiment 3: a pronounced, stable interaction indicative of shape-color binding. Because color was again an irrelevant dimension and because the colors were not pre-experimentally related to the shapes of the stimuli, this finding supports the idea that the *objecthood* of stimuli changes the way they are processed and the way their features are integrated. Once again, this suggests that higher-level representations are involved in integrating real-world shapes with color features. Indeed, it is plausible to assume that long-term object representations are less context-specific than situational models of conjunction probabilities of arbitrary features and, hence, less selective with respect to the feature dimensions related to the object at hand. In other words, although one may learn to neglect the color of shape-color conjunctions if the task at hand only requires attending to shape, it makes little sense to drop color information as a constituent of the long-term representation of an apple only because in one given situation the color of a particular apple played no role. We will elaborate on the architectural implications of this consideration in the General Discussion.

General Discussion

The four experiments of this study aimed at addressing the relation between short-term binding effects and long-term learning: to what extent are binding and learning processes independent, to what extent do they interact? We considered short-term binding aftereffects and learning in terms of both experimentally induced feature-conjunction biases and stably learned natural feature conjunctions. Although it is clear that more research on this issue is necessary, we take our findings to point to an independence of binding and learning, at least with respect to direct interactions. That is, we think it is justified to reject the strong-dependence hypothesis outlined in the introduction. The different result patterns we obtained in Experiments 1 and 2 on the one hand and in Experiments 3 and 4 on the other further suggest that it matters whether binding is restricted to arbitrary, frequently changing visual features or whether overlearned objects are involved. In our view, explaining the outcome of this study requires the consideration of three different systems or representational levels: a low-level representation of features in feature maps; a higher-level, long-term representation of objects; and a working memory system in which situational contingencies are temporarily stored. The relationship between the first two of these systems is captured in Figure 7.

On the one hand, we assume that the binding process proper is automatic and not directly impacted by higher-level signals from either inferotemporal or prefrontal cortex. On the other hand,
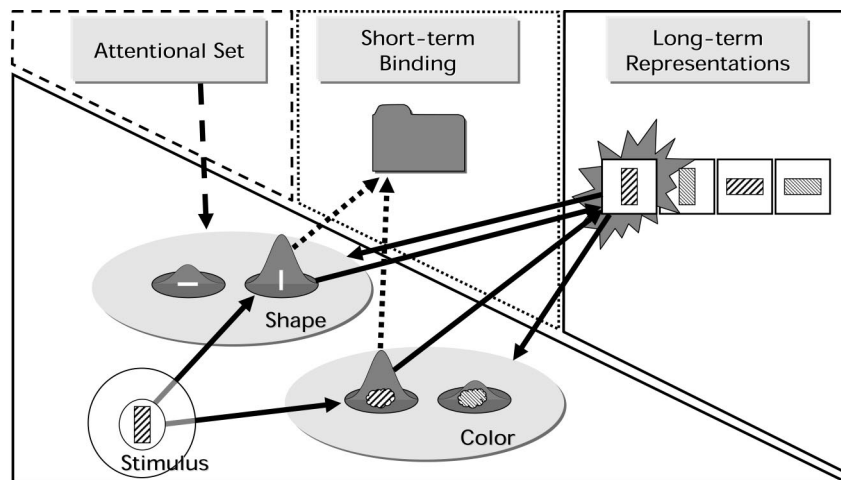

Figure 7. Sketch of the systems and processes involved in feature integration. In the example, a vertical bar of a particular color (indicated by pattern) is processed. It activates its feature codes in the shape (or orientation) map and the color map. The degree to which processing a feature activates its code is modulated by the current attentional set, which provides top-down support for features coded on task-relevant feature dimensions. In the present study, shape was relevant but color was not; therefore, the attentional set would be biased toward shapes a practice increases, that is, attention becomes increasingly selective. All feature codes whose activations exceed a particular integration threshold are bound, that is, temporarily linked to an "event file" (Hommel, 2004). After some practice in our current task, this would be true for shapes but less so for colors. However, overlearned feature combinations, as well as single features belonging to overlearned combinations, are concurrently registered by conjunction detectors, that is, object representations in long-term memory. Once a conjunctive code is activated it "backward-primes" the feature dimensions on which features belonging to the represented object or conjunction are coded. This can overrule or counteract the impact of the attentional set and thereby lead to the inclusion of a feature coded on a task-irrelevant dimension. Note that the mechanism that produces short-term facilitation of frequent conjunctions is considered to act upon processes that follow feature integration, which is why it is not included in this figure.

however, which features are considered for binding does depend on attentional settings (Hommel, 2004), which again are affected by situational models in working memory and, if available, object representations in long-term memory. In the following, we will elaborate this theoretical framework and consider its neural plausibility.

Higher-Order Object Representation and Low-Level Binding

Experiments 1 and 2 provide evidence that processing the shape (or orientation) and color of simple geometric stimuli leads to a binding of the codes representing the features. However, in contrast to the overlearned objects used in Experiments 3 and 4, the geometric stimuli produced binding over a couple of trials only. This means that the presence or absence of long-term representations of the particular stimuli makes a difference for binding. On the other hand, however, we found no evidence for any direct interaction between short-term or long-term object representations and binding. How can this be explained? We think the key to understanding this somewhat-complex relationship between learning and binding requires the distinction between at least two levels of representation and integration: a lower representational level at which features are temporarily linked (see features maps in Figure 7, and the event file to which the codes relate) and a higher representational level at which integrated feature assemblies are stored (see long-term representations). Apparently, linking features at the proposed lower level does not *directly* translate into having stored feature links at the higher level, and having a stored feature link at the higher level does not *directly* impact creating a link at the lower level.

Feature-binding problems often are discussed with respect to the visual system, where the strong evidence for a whole multitude of feature maps (Felleman & van Essen, 1991; Lamme & Roelfsema, 2000) makes the need for integration processes particularly obvious. As discussed in the introduction, a promising candidate mechanism for fast binding of low-level visual features like orientation or color, is the synchronization of neural responses (Gray, König, Engel, & Singer, 1989; von der Malsburg, 1999). As compared with other mechanisms, synchronization would not only be a fast and flexible mechanism but would also enable the representation of a very large number of however novel and arbitrary feature combinations. Indeed, there is significant evidence in support of the idea that synchronization in the gamma frequency band (high frequency EEG activity above 30 Hz) plays a role in visual feature binding (Engel & Singer, 2001), visual working memory (Luck & Vogel, 1997; Raffone & Wolters, 2001), and consciousness (Engel & Singer, 2001).

However, the visual system does not only contain of low-level feature maps with a high spatial resolution such as in V1 and V2/V3. Higher level neurons coding for more complex shapes and multifeature objects can be found at later stages in the occipito-temporal (ventral) stream, like in V4 and IT. Converging feed-forward connections are likely to enable increased response selectivity and transmission of signals for fast bottom-up processing, and feedback diverging synapses to mediate attentional and learning-based modulation of neural responses. Numerous studies provide evidence that convergence plays a more important role at this higher representational level. For instance, some cells have

been shown to be selective to stimuli as complex as faces (Young & Yamane, 1992) or, in the posterior inferior temporal cortex, cells that are selective for conjunctions of a striped patch and flanking black spots (Tanaka, Saito, Fukada, & Moriya, 1991). Hence, even if we exclude convergence as the only integration mechanism, there are good reasons to believe that at least some feature conjunctions are encoded by assemblies of a limited number of selectively tuned neurons in inferotemporal cortex, which are adapted and shaped by Hebbian learning (e.g., Amit, 1995).

Thus, we suggest that temporary binding (e.g., by synchronization) and coding by convergence are not as exclusive as previously held (e.g., Jellema & Perrett, 2002; Singer, 1994) but, rather, may coexist to solve binding problems at different levels (Hommel, 2004; Singer, 1999). In particular, the recognition of familiar objects is achieved by assemblies of highly selective conjunction detectors that only emerge for behaviorally relevant, frequently occurring events and that change only slowly through (probably Hebbian) learning. In contrast, frequently changing or novel combinations of arbitrary visual features are coded by synchronizing relatively raw feature codes represented at feature maps.

The Role of Frequency Learning and Attentional Weighting

Let us now consider how low-level feature binding may be affected by the objecthood of the stimuli involved and the probability of particular feature combinations. As we have pointed out, we found no evidence for any direct impact of frequency or familiarity manipulations on partial-overlap costs, our measure of feature binding. And yet, more probable combinations were processed faster than less probable combinations. This means that the system was biased toward more likely feature conjunctions but that this bias affected processing only after features were bound or, in the case of partial overlap, rebound. One possibility is that the shape-color associations underlying this bias represent a first step in the emergence of a new object representation. However, this idea does not seem to fit with the lack of transfer of the frequency bias to the test phase and to the different result patterns in Experiments 1 and 2 on the one hand and Experiments 3 and 4 on the other. Alternatively, frequency-based expectations may be incorporated into situational models held in working memory. For instance, low-level integration may run autonomously but its outcome may be registered and processed more quickly if it fits situation-specific expectations. In any case, however, it is important to note that the speed of what we attribute to low-level feature integration is unaffected by top-down expectations.

And yet, there was evidence for an indirect modulation of low-level integration by top-down processes. When we used stimuli made up of arbitrary feature conjunctions (combinations of shape and color) we found our measure of feature binding to be rather instable and it even disappeared over time, whereas lifelong practiced stimuli yielded robust and stable effects. We attribute the first finding to adaptive feature weighting (Hommel, Müsseler, Aschersleben & Prinz, 2001), that is, to the dynamic weighting of feature dimensions according to their contribution to task performance. Feature weighting is an attentional process that selectively prepares the cognitive system for the differential processing of relevant (i.e., to-be-attended) and irrelevant (i.e., to-be-ignored) features of anticipated perceptual events (cf., Bundesen, 1990;

Found & Müller, 1996). Because the color of stimuli is rather salient and likely to be helpful in discriminating targets from nontarget stimuli, such as the fixation point or cues, the weight of the color dimension is unlikely to be zero, at least at the beginning of an experiment. With increasing practice, however, people will fine-tune the weights of the perceptual dimensions to better reflect their use for current performance. As color was irrelevant to the task, this is likely to have led to a continuous decrease of the color-dimension weight. (Note that the weak contribution from color is not due to a particularity of this dimension, as the contribution from shape is as weak if S2 is not defined by shape: Hommel, 1998.) If we further assume that the weight of a perceptual dimension determines the probability that the corresponding features are considered in perceptual binding (Hommel, 2004), it is easy to see why the decrease in color weighting eliminated partial-overlap costs: the color feature was activated to degree that was insufficient for binding, so that its code was not involved anymore.

Using real-life objects brought long-term object representations into play. These representations must have emerged from numerous encounters with the represented objects, which implies that they do not include situational particularities, such as the task-specific value of one or the other feature dimension. Accordingly, it makes sense to assume that the involvement of long-term object representations top-down primed all of the feature dimensions defining the object to an equal degree. Indeed, there is considerable evidence that processing one feature of an object automatically opens the attentional gate to other features of this object (Baylis & Driver, 1992; Duncan, 1984; Kahneman & Henik, 1981). This top-down priming effect may prevent or overrule the practice-induced unweighting of nominally irrelevant feature dimensions, and thereby keep the contribution of features defined on such a dimension sufficiently strong to stay involved in binding. Our present findings might be interpreted to mean that (pictures of) real objects have a reliable top-down impact (Experiments 3 and 4) whereas frequent combinations of arbitrarily chosen features do not (Experiments 1 and 2). However, note that there were several hints to a reduction of practice-induced inattention to color for frequent feature combinations. The fact that these hints failed to reach statistical significance may be simply due to insufficient learning, suggesting that more repetitions may make even previously arbitrary feature combinations to be represented as, and to act like "real" objects.

Levels of Binding and Learning

Taken altogether, our experiments suggest the existence of different levels of integration and conjunction learning in the human brain. A first level flexibly combines entries in low-level feature maps (such as in V1, V2, and V4). Binding at this level is highly context-sensitive, suggesting that features are linked to task or context information (cf., Waszak, Hommel, & Allport, 2003).

A second type of short-term learning expresses itself in rather local contingencies, such as the probability of particular feature conjunctions. This type of learning is also transient and task specific, which may point to the involvement of situational models in working memory. It may lead to a faster readout of expected feature conjunctions and/or the lowering of thresholds in *responding* to expected conjunctions. In any case, however, frequency-based expectations do not seem to speed up (re-)binding or access

to object representations but seem to act on subsequent processing steps.

Third, known, overlearned objects are represented at a more integrated level, presumably by means of conjunction detectors. At this level, familiar objects are long-term encoded, and activating their codes provides top-down priming of object-related feature dimensions. In particular, coding familiar objects has the consequence that activating one part of an object representation spreads to the whole integrated assembly, in a kind of pattern completion process (Hommel, 2004). This makes it difficult to isolate the contributions of individual components (e.g., feature codes) in the assembly, which among other things overrules possibly differential attentional weights for object features provided by the current attentional set. Experiment 4 suggests that these assemblies need not be restricted to overlearned feature values but may also be updated by and "capture" context-specific features, like colors.

The proposed architecture leads to interesting empirical predictions with respect to quite a number of tasks and phenomena. For instance, consider the observation that people are able to hold multiple feature conjunctions in visual working memory (Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001; but see Wheeler & Treisman, 2002). Our architecture would predict that conjunctions between ecologically unrelated features (such as the commonly chosen geometric shapes and colors) are more difficult to hold and less robust against interference and decay than conjunctions involving the shapes of real-world objects (like the bananas and strawberries we used). Hence, the present estimations based on conjunctions between arbitrary features may underestimate the true capacity of human visual working memory.

To conclude, our findings suggest that the representation of feature conjunctions is a multicomponent process involving several time scales and levels of integration. They also suggest that the interaction between top-down attentional processes and automatic binding processes is dynamic and adaptive to task constraints. It remains to be seen how integrational structures or event files of different nature behave over short and long time scales with intentional maintenance of binding codes in working memory. These representational behaviors are likely to depend upon the interactions between prefrontal cortex, posterior cortical cortices, and premotor cortex and perhaps involve long-range neural synchronization (Gross, Schmitz, Schnitzler, Kessler, Shaprio, Hommel, & Schnitzler, 2004; Tononi, Sporns, & Edelman, 1992). Related neuroimaging and neurophysiological investigations, as well as large-scale neurocomputational modeling, will play a crucial role in answering these core neurocognitive questions.

References

Amit, D. J. (1995). The Hebbian paradigm reintegrated: Local reverberations as internal representations. *Behavioral and Brain Sciences, 18,* 617–626.

Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology. *Perception, 1,* 371–394.

Baylis, G. C., & Driver, J. (1992). Visual parsing and response competition: The effect of grouping factors. *Perception & Psychophysics, 51,* 145–162.

Braitenberg, V., & Schüz, A. (1991). *Anatomy of the cortex: Statistics and geometry.* Heidelberg, Germany: Springer.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97,* 523–547.

Cowey, A. (1985). Aspects of cortical organization related to selective attention and selective impairments of visual perception: A tutorial review. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 41–62). Hillsdale, NJ: Erlbaum.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113,* 501–517.

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience, 2,* 820–829.

Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences, 5,* 16–25.

Fell, J., Fernandez, G., Klaver, P., Elger, C. E., & Fries, P. (2003). Is synchronized neuronal gamma activity relevant for selective attention? *Brain Research Reviews, 42,* 265–272.

Felleman, D. J., & van Essen, D. C. V. (1991). Distributed hierarchical processing in the primate visual cortex. *Cerebral Cortex, 1,* 1–47.

Found, A., & Müller, H. J. (1996). Searching for features across dimensions: Evidence for a dimensional weighting account. *Perception & Psychophysics, 58,* 88–101.

Gordon, R. D., & Irwin, D. E. (1996). What's in an object file? Evidence from priming studies. *Perception & Psychophysics, 58,* 1260–1277.

Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature, 338,* 334–337.

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., & Schnitzler, A. (2004). Long-range neural synchrony predicts temporal limitations of visual attention in humans. *Proceedings of the National Academy of Sciences, USA, 101,* 13050–13055.

Hebb, D. O. (1949). *The organization of behavior.* New York: Wiley.

Henderson, J. M. (1994). Two representational system in dynamic visual identification. *Journal of Experimental Psychology: General, 123,* 410–426.

Henderson, J. M., & Anes, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 826–839.

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus response episodes. *Visual Cognition, 5,* 183–216.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8,* 494–500.

Hommel, B., & Colzato, L. S. (2004). Visual attention and the temporal dynamics of feature integration. *Visual Cognition, 11,* 483–521.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24,* 849–878.

Jellema, T., & Perrett, D. I. (2002). Coding of visible and hidden objects. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and performance XIX* (pp. 356–380). Oxford, England: Oxford University Press.

Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 181–211). Hillsdale, NJ: Erlbaum.

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

Lamme, V. A., Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23,* 571–579.

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

Miltner, W. H., Braun, C., Arnold, M., Witte, H., & Taub, E. (1999). Coherence of gamma-band EEG activity as a basis for associative learning. *Nature, 397,* 434–436.

Mondor, T. A., Hurlburt, J., & Thorne, L. (2003). Categorizing sounds by pitch: Effects of stimulus similarity and response repetition. *Perception & Psychophysics, 65,* 107–114.

Park, J., & Kanwisher, N. (1994). Negative priming for spatial locations: Identity mismatching, not distractor inhibition. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 613–623.

Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience, 13,* 766–785.

Singer, W. (1994). The organization of sensory motor representations in the neocortex: A hypothesis based on temporal coding. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 77–107). Cambridge, MA: MIT Press.

Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron, 24,* 49–65.

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.

Tononi, G., Sporns, O., & Edelman, G. M. (1992). Reentry and the problem of integration of multiple cortical areas: Simulation of dynamic integration in the visual system. *Cerebral Cortex, 2,* 310–335.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6,* 171–178.

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

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 and Performance, 27,* 92–114.

von der Malsburg, C. (1999). The what and why of binding: The modeler's perspective. *Neuron, 24,* 95–104.

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology, 46,* 361–413.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131,* 48–64.

Young, M. P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science, 256,* 1327–1331.