

---

# Units of visual individuation in rhesus macaques: objects or unbound features?

---

Erik W Cheries, George E Newman, Laurie R Santos, Brian J Scholl

Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205, USA;  
e-mail: Erik.Cheries@yale.edu

Received 27 May 2005, in revised form 8 November 2005; published online 9 June 2006

---

**Abstract.** Vision begins with the processing of unbound visual features, which must eventually be bound together into object representations. Such feature binding is required for coherent visual perception, and accordingly has received a considerable amount of study in several domains. Neurophysiological work, often in monkeys, has revealed the details of how and where feature binding occurs in the brain, but methodological limitations have not allowed this research to elucidate just how feature binding operates spontaneously in real-world situations. In contrast, behavioral work with human infants has demonstrated how we use simpler unbound features to individuate and identify objects over time and occlusion in many types of events, but this work has not typically been able to isolate the role of feature binding in such processing. Here we provide a method for assessing the spontaneity and fidelity of feature binding in non-human primates, as this process is utilized in real-world situations, including simple foraging behaviors. Using both looking-time and manual-search measures in a natural environment, we show that free-ranging rhesus macaques (*Macaca mulatta*) spontaneously bind features in order to individuate objects across time and occlusion in dynamic events. This pattern of results demonstrates that feature binding is used in subtle ways to guide ecologically relevant behavior in a non-human animal, spontaneously and reliably, in its natural environment.

## 1 Introduction

Many aspects of visual perception are, at root, attempts to grapple with *inverse problems*. Objects in the world have numerous properties, many of which are represented in a fragmented and compressed form in the light they selectively reflect and absorb. The visual system then receives this light as input, and attempts to invert the function, so as to arrive at an interpretation of a scene which accords with actual objects and properties in the world. Such inverse problems are especially salient in the context of individual visual features. In the computation of depth from disparity, for example, objects in the world initially possess 3-D shape and location, which is collapsed in our visual input into a pair of 2-D images (each initially processed in distinct neurophysiological loci). The visual system must then put this information ‘back together’ in such a way as to recover the missing third dimension. Depth is in no way a special case in this respect; similar types of inverse problems occur with several other individual visual features, such as the perception of lightness (in which illumination must be entangled from reflectance, despite being combined in the reflected light).

A related inverse problem occurs in processing which must span multiple visual features. Objects in the world possess multiple properties: one object might be both yellow and round, while another is green and triangular. This ‘binding’ of features, however, is not explicit in reflected light, but must be reconstructed. This ‘binding problem’ becomes particularly acute in visual processing, because different features are initially processed via functionally and neurophysiologically distinct operations. As a result, features must later be (re-)bound into discrete object representations, which make explicit which features ‘go together’. This is the so-called *binding problem* in visual perception.

### 1.1 *Research on feature binding*

Solving the binding problem is clearly critical for coherent perception of the visual world, and accordingly has been the focus of a vast amount of research across several levels of vision science [for overviews, see Müller et al (2001), Treisman (1999); for a model in which the binding problem may simply not arise, though, see Riesenhuber and Poggio (1999)]. This work has collectively attempted to determine how (and where, and when) the visual system processes fragmented individual visual features and combines them to produce our phenomenologically unified percepts. The nature of such processing has been illuminated in several ways: by neurophysiological research on non-human animals, psychophysical experiments in human adults, and looking-time studies in human infants. We describe each of these research programs in turn.

Numerous neurophysiological studies have directly suggested that rhesus monkeys perceive bound features, by observing brain areas and individual neurons that selectively respond only to moderately complex conjunctions of features rather than to individual properties in isolation (eg Tanaka 1993, 1996; Tanaka et al 1991). Earlier neurophysiological studies, moreover, highlighted the existence of the binding problem in physical terms by demonstrating that individual properties of objects (eg color, motion, shape) are analyzed at least partially independently and are processed in separate, specialized neural circuits (eg Desimone et al 1985; Livingstone and Hubel 1984, 1987, 1988; cf DeYoe and Van Essen 1988). Casting the binding problem at this level also illustrates why shared spatial locations do not dismiss the problem: though the features which must be bound typically do arise from the same location in the visual field, many neurons in the visual system have large receptive fields—often spanning several degrees (eg in V4) or even the majority of the central visual field (eg in IT; Desimone et al 1984; Desimone and Schein 1987; Gattass et al 1988). More recent work in this tradition has focused on the details of how and when the binding process itself occurs. For example, neurophysiological studies with rhesus macaques (*Macaca mulatta*) in carefully controlled laboratory settings have revealed that the synchrony (eg Engel and Singer 2001; Singer and Gray 1995; cf Thiele and Stoner 2003) and latency (eg Gawne et al 1996) of neuronal activation may constitute a mechanism for binding information in the visual field. Similarly, other theories posit that selective attention acts on a neurophysiological level to shrink the spatial extent of receptive fields (eg Lucks et al 1997; Moran and Desimone 1985; Reynolds et al 1999), thus increasing the plausibility of a solution to the binding problem based on shared spatial locations.

Psychophysical work with human adults has also helped to reveal how and when feature binding occurs, and under what conditions it breaks down. Although features such as color and shape normally appear inextricably bound into objects in everyday perceptual experience, certain experimental manipulations can frustrate such processing, yielding ‘illusory conjunctions’ in which features from distinct objects are mistakenly bound together (Treisman and Schmidt 1982). Perhaps the most salient demonstration of illusory conjunctions is in certain rare neuropsychological disorders. In Balint syndrome, for example, patients (typically with bilateral parietal damage) can often see only a single object at a time, but certain visual features present in a scene can become unbound and may perceptually float through the single visible object. Thus, if such a patient is presented with an array of colored geometric shapes, they may perceive only a single square, but the perceived color of that square may then shift over time, from red to blue to green, as other colors in the scene become mistakenly and transiently bound to it (eg Friedman-Hill et al 1995; Robertson et al 1997; see also Humphreys et al 2000).

Illusory conjunctions can also occur in normal subjects, such that a green triangle and a yellow ring might be perceived as a yellow triangle and a green ring. Such phenomena typically require brief masked displays to impair feature binding, but some illusory

conjunctions—eg between color and motion—can persist indefinitely in perceptual experience with free viewing even when observers know about the correct binding (Wu et al 2004). The mechanics of feature binding have been studied by exploring the stimulus factors that affect the likelihood and extremity of illusory conjunctions. For example, illusory conjunctions are more severe: for spatially ambiguous features, eg those in the periphery (Prinzmetal et al 1995); with certain visual features (such as topological closure) which are thought to be processed especially early (Chen and Zhou 1997); between feature clusters that are already perceptually grouped by spatial factors (eg Cohen and Ivry 1989); between feature clusters that are both parsed as figure, compared to features that span both figure and ground (Treisman 1988); etc. These factors can in turn be interpreted as helping to determine just how and when binding occurs.

Of course, in most situations successful feature binding does occur, and indeed is often perceptually irresistible: in many situations, for example, it is not possible to attend to one feature of an object without thereby automatically attending to the other features of that object (Kahneman and Henik 1981; O'Craven et al 1999). Moreover, many other visual and cognitive processes appear to operate over object representations into which features are bound: for example, visual working memory may in some cases be object-based rather than feature-based, such that featurally complex scenes (in which objects differ in color, texture, shape, and orientation) can be encoded just as easily and quickly as featurally sparse scenes [eg containing identical shapes differing only in their color; eg Luck and Vogel (1997); cf Alvarez and Cavanagh (2004)].

A third (and largely independent) avenue for exploring the nature of feature binding comes from work with human infants, who do not always employ the features of an object in the same manner as adults. Much recent work has explored the ways in which various features and their combinations are and are not used in simple individuation tasks, which test infants' ability to represent objects as numerically distinct individuals. Before 12 months of age infants perceive (and habituate to) feature contrasts (Needham 1998; Wilcox 1999), but do not use these contrasts in some situations to individuate and track objects over time. This has been shown by using both looking-time (Tremoulet et al 2000; Xu and Carey 1996) and manual-search measures (Van de Walle et al 2000). If two objects with very different visual features—for example, a duck and a truck—sequentially appear from behind a screen, only infants older than 12 months will typically infer that there must be two distinct individuals present behind the screen, and will thus look longer when a screen drops to reveal only a single object [eg Xu and Carey (1996); see Tremoulet et al (2000) for similar experiments with even simpler featural contrasts]. Similarly, if a duck and a truck are sequentially removed and then replaced from a box, and infants are then allowed to reach into the box, only infants older than 12 months will typically continue to reach for a second object, after having retrieved only a single object (Van de Walle et al 2000).

Moreover, related studies have demonstrated that features are not always bound in infants' representations of such situations, since different features 'come online' to be used in such processes at different times—for example, with shape and size being used several months before pattern information, which in turn is used several months before color (Wilcox 1999; though cf Kaldy et al 2004). Other infant research has looked in more detail at feature binding per se. For example, infants as young as 1 month of age will dishabituate to situations in which a green triangle and a yellow ring emerge from a momentary period of occlusion as a yellow triangle and a green ring—maintaining all of the same features, changing only their bindings (eg Bushnell and Roder 1985; Kaldy and Leslie 2003; Slater et al 1991; Taga et al 2002).

---

## 1.2 *A missing link?*

In the experiments reported below, we adapt various tools and questions from across these disparate areas of research, to ask questions about how feature binding is used in the spontaneous behavior of non-human primates. Neurophysiological research on feature binding in non-human primates has begun to clarify just how feature binding processes are realized in the brain, and in what situations monkeys are able to bind features into coherent object representations. However, this work is typically divorced from ecological constraints, and the behavioral components of such experiments typically require substantial training periods in order to ensure that subjects perform the task. As a result, we cannot draw any straightforward conclusions from this research about the fidelity and spontaneity of feature binding in real-world situations, or how feature binding may or may not influence natural behaviors of monkeys.

Psychophysical work with human adults is in a similar situation. On the one hand, this research program has revealed a detailed picture of how and when feature binding operates and breaks down. However, methodological limitations of this work have been especially well suited to asking about *whether* feature binding occurs, without being able to determine just how feature binding is used in online ecological behavior. This contrast is highlighted by some of the infant work discussed above, since in some cases we know that infants perceive various features without also using them in the service of more complex tasks. Thus, simple psychophysical measures may overestimate the fidelity of feature binding as it is actually used in the real world. A second limitation of the adult work, of course, is that it is limited to verbally fluent populations. Thus, if we want to discover how feature binding operates in non-human animals, we must look to other methods.

Finally, research with human infants has devised clear paradigms for studying feature binding in nonverbal populations, involving both implicit looking-time measures and active search measures. However, this research has never examined just how (or whether!) feature binding will guide infants' spontaneous behaviors. On one hand, studies that have directly explored feature binding in infancy have not tied this ability to object individuation and tracking: when infants dishabituate to a mismatched feature pair, they may not necessarily recognize that the objects have changed or that new objects are present—but may instead simply be evincing surprise that a novel feature combination is now present. On the other hand, other looking and reaching experiments have the ability to directly assess infants' object-individuation abilities—determining, for example, whether a novel feature will lead infants to conclude that another object must still be present behind a screen. However, work with these paradigms to date has been limited to the study of individual visual features, and was never extended to 'feature miscombinations' in order to address feature binding.

In the present study, we thus attempt to exploit the best aspects of each of these research areas. As criteria for this study, we thus require: (i) the use of non-human primates as subjects; (ii) an experimental manipulation that can isolate the role of feature binding per se; (iii) a paradigm that can test how feature binding may or may not be *used* in the service of object individuation, beyond simple sensitivity to feature combinations; (iv) the ability to test how feature binding may or may not operate in real-world situations, beyond the laboratory; and (v) the ability to test the fidelity of feature binding in *spontaneous* performance, without any previous training.

## 1.3 *The current experiments*

In the experiments reported below, we attempt to satisfy these criteria by adapting paradigms used in infant work for use with non-human primates in natural foraging situations, in order to test for how feature binding is spontaneously used in real-world perception. In experiment 1 we employ a looking-time method in order to determine the

fidelity with which free-ranging rhesus macaques will spontaneously bind features when passively observing displays with simple geometric objects. In experiment 2, we employ an active manual-search measure with food objects, to explore the degree to which feature binding is actually used in object individuation during real-world foraging behavior.

## 2 Experiment 1: Looking time

Using a looking-time method, we first familiarized rhesus monkeys with two geometric objects that could differ in their colors and/or shapes. After repeatedly viewing these objects, the monkeys were then presented with test displays in which the objects: (i) remained identical; (ii) changed their colors (eg two green shapes both turning yellow); (iii) changed their shapes (eg two rings both turning into triangles); (iv) swapped locations; or (v) swapped features (eg a yellow ring and green triangle turning into a green ring and yellow triangle). We examined how long the monkeys looked at these test events, as a measure of the degree to which they discriminated these types of changes. Critically, the final ‘feature swap’ condition did not introduce any novel features into the display as a whole, so that successful discrimination would have to be based on bound object representations rather than on an analysis of any individual features.

### 2.1 Method

2.1.1 *Subjects.* We tested free-ranging adult rhesus macaques at the Cayo Santiago field station (Rawlins and Kessler 1987), a population of 800 individuals (readily identifiable via ear notches and chest tattoos), all well habituated to human experimenters. Similar paradigms have been used with this population in a number of previous studies (eg Flombaum et al 2005; Hauser et al 1996; Munakata et al 2001; Santos and Hauser 2002; Santos et al 2003). This experiment involved several conditions as described below. A total of 111 monkeys were tested. As described below, another 96 monkeys were approached but not fully tested, owing to interference from other monkeys, inattention to the display, lack of approach (in the case of experiment 2), or because they had been tested previously.

2.1.2 *Apparatus.* Subjects were familiarized to a display of two geometric objects presented on a white foamcore stage. Each object was shaped either as a ring or a triangle (roughly 6 cm by 6 cm), and was constructed from either bright yellow or bright green Play-Doh<sup>®</sup> (see figure 1). The foamcore stage consisted of a base (a 24 cm × 10 cm back, and a 24 cm × 20 cm floor), and a screen (24 cm × 8 cm).



**Figure 1.** A photograph of two of the four objects used in experiment 1.

**2.1.3 Procedure.** As in previous looking-time studies with this population, subjects were chosen opportunistically by searching out individuals who were alone and in a seated position. Two experimenters ran each session. The first experimenter (the ‘presenter’) sat approximately 1.5 m in front of the subject and performed all of the actions on the apparatus. The second experimenter (the ‘cameraperson’) stood directly behind the first and recorded the session with a Sony Hi-8 Handicam. The cameraperson stood such that the apparatus was out of view and therefore was able to record the session while remaining blind to the experimental condition.

Previous experiments (see Santos et al 2003) have demonstrated that subjects in this population typically habituate (ie show reliably less looking) after approximately three familiarization trials. We therefore presented each subject with three familiarization trials followed by a single test trial. In the familiarization trials, the presenter lifted the screen of the box to reveal two objects (as described below) resting on the stage (see figure 1). After lifting the screen, the experimenter called out “now” and the subject’s looking time was recorded for the next 10 s. After the 10 s, the presenter replaced the screen, removed the objects (out of the subject’s view), added the objects for the next trial, reacquired the monkeys’ attention, then began the next trial. (The objects were always replaced on every trial—both familiarization and test trials—even in subsequent trials employing the identical objects.)

Videotapes were later recorded onto an iBook laptop computer and were analyzed with Adobe Premiere software. A single experimenter who was blind to the test condition coded each subject’s looking during each 33 ms frame of the 10 s looking period that followed each test trial. Looking was defined as periods in which the subject’s head and eyes were both oriented toward the display. One quarter of the trials in each condition were independently scored by a second coder; the two coders were highly reliable ( $r = 0.97$ ), and the first coder’s judgments were used in case of a disagreement.

**2.1.4 Conditions.** We tested several different conditions, as follows, each of which used a different group of subjects. In the Whole Object Change condition ( $n = 14$ ; figure 2), subjects were always familiarized with two identical objects (eg two yellow triangles), both of which changed their shape and color at test (eg to two green rings). The Whole Object Control condition ( $n = 11$ ) involved the identical familiarization events, and a final test trial which involved no changes. We attempted to test 29 additional subjects split between these two conditions, but these individuals were dropped from the analysis owing to either previous testing, interference from other subjects, or disinterest.

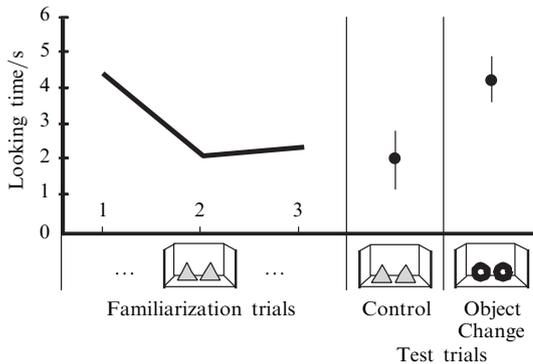
In the Shape Change condition ( $n = 12$ ; figure 3a), subjects were always familiarized with two objects with the same shape but different colors (eg green and yellow rings), which changed their shapes but not colors at test (eg to green and yellow triangles). The Shape Control condition ( $n = 8$ ) involved the identical familiarization events, and a final test trial which involved no changes. In the Color Change condition ( $n = 17$ ; figure 3b), subjects were always familiarized with two objects with the same color but different shapes (eg a green ring and triangle), which changed their colors but not shapes at test (eg to a yellow ring and triangle). The Color Control condition ( $n = 7$ ) involved the identical familiarization events, and a final test trial which involved no changes. We attempted to test 49 additional subjects split between these four conditions, but these individuals were dropped from the analysis owing to either previous testing, interference from other subjects, or disinterest.

In the Feature Swap condition ( $n = 17$ ; figure 4), subjects were always familiarized with two objects which differed in both their colors and shapes (eg a green ring and a yellow triangle), and these features were swapped for the test trial (eg to a yellow ring and a green triangle). The Feature Swap Control condition ( $n = 12$ ) involved the identical familiarization events, and a final test trial which involved no changes.

Finally, an additional Position Change Control condition ( $n = 13$ ) involved the identical familiarization events, but at test these same two objects had simply switched relative positions on the stage. We attempted to test 18 additional subjects split between these three conditions, but these individuals were dropped from the analysis owing to either previous testing, interference from other subjects, or disinterest.

## 2.2 Results

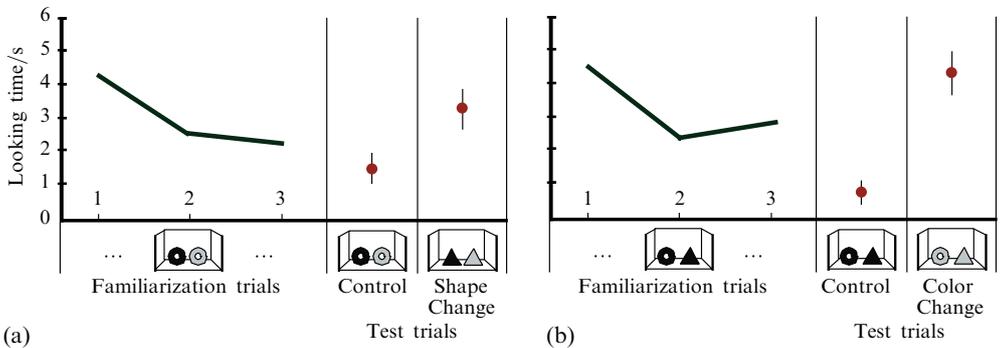
**2.2.1 Whole Object Change conditions.** As is clear from inspection of figure 2, subjects habituated during the initial familiarization trials, and then recovered looking time in Whole Object Change test trials, but not in Whole Object Control test trials. These impressions were verified by the following statistical analyses. During familiarization, subjects looked reliably less in the third familiarization trial than in the first (paired  $t$ -test:  $t_{24} = 3.59$ ,  $p < 0.01$ ). There was a reliable difference between duration of looking in the test trials (unpaired  $t$ -test:  $t_{23} = 2.73$ ,  $p = 0.01$ ): subjects in the Whole Object Change group looked significantly longer than subjects in the baseline Whole Object Control group. This comparison was also significant via a non-parametric test (Mann–Whitney:  $z = 2.90$ ,  $p > 0.01$ ). The two test groups also differed in their degree of dishabituation at test: subjects in the Whole Object Change group recovered looking time relative to the final familiarization trial (paired  $t$ -test:  $t_{13} = 2.31$ ,  $p = 0.04$ ), but those in the Whole Object Control group did not ( $t_{10} = 0.12$ ,  $p = 0.90$ ). Ten of the thirteen subjects in the Whole Object Change test group dishabituated in this manner (Wilcoxon signed rank:  $z = 2.13$ ,  $p = 0.03$ ).



**Figure 2.** Sample stimuli and results from the Whole Object Change conditions of experiment 1. The graph depicts mean looking time  $\pm$  standard error for the familiarization and test trials.

**2.2.2 Shape Change conditions.** As is clear from inspection of figure 3a, subjects habituated during the initial familiarization trials, and then recovered looking time in the Shape Change test trial, but not in the Shape Control test trial. These impressions were verified by the following statistical analyses. During familiarization, subjects looked reliably less in the third familiarization trial than in the first (paired  $t$ -test:  $t_{19} = 4.17$ ,  $p < 0.01$ ). There was a marginal difference between duration of looking in the test trials (unpaired  $t$ -test:  $t_{18} = 2.01$ ,  $p = 0.059$ ): subjects in the Shape Change group looked longer than subjects in the baseline Shape Control group. This comparison was significant via a non-parametric test (Mann–Whitney:  $z = 2.04$ ,  $p > 0.04$ ). The two test groups also differed in their degree of dishabituation at test: subjects in the Shape Change group recovered looking time relative to the final familiarization trial (paired  $t$ -test:  $t_{11} = 2.16$ ,  $p = 0.05$ ), but those in the Shape Control group did not ( $t_7 = 1.31$ ,  $p = 0.23$ ). Ten of the twelve subjects in the Shape Change test group dishabituated in this manner (Wilcoxon signed rank:  $z = 2.35$ ,  $p = 0.02$ ).

**2.2.3 Color Change conditions.** As is clear from inspection of figure 3b, subjects habituated during the initial familiarization trials, and successfully discriminated the two test displays by looking longer in the Color Change test trials, and less in the Color Control

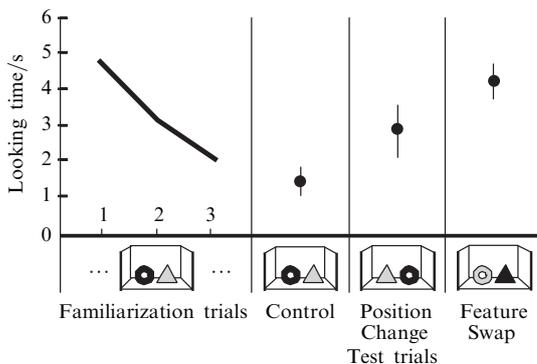


**Figure 3.** Sample stimuli and results from (a) the Shape Change conditions and (b) the Color Change conditions of experiment 1. The graphs depict mean looking time  $\pm$  standard error for the familiarization and test trials.

test trials. These impressions were verified by the following statistical analyses. During familiarization, subjects looked reliably less in the third familiarization trial than in the first (paired  $t$ -test:  $t_{23} = 2.68$ ,  $p < 0.01$ ). There was a reliable difference between duration of looking in the test trials (unpaired  $t$ -test:  $t_{22} = 3.07$ ,  $p < 0.01$ ): subjects in the Color Change group looked significantly longer than subjects in the baseline Color Control group. This comparison was also significant via a non-parametric test (Mann–Whitney:  $z = 2.89$ ,  $p < 0.01$ ). The two test groups also differed in their degree of continued habituation at test: subjects in the Color Change group did not show a difference in looking time between the final familiarization and test trial (paired  $t$ -test:  $t_{16} = 1.18$ ,  $p = 0.26$ ), but those in the Color Control group showed a marginal trend to look less than in the last familiarization trial ( $t_6 = 2.09$ ,  $p = 0.08$ ). Non-parametrically, six of the seven subjects in the Color Control group looked less in this manner (Wilcoxon signed rank:  $z = 2.20$ ,  $p = 0.03$ ).

Across both the Shape Change and Color Change conditions, there was no reliable difference in looking time between the Color Change and Shape Change test groups, either parametrically ( $t_{27} = 0.68$ ,  $p = 0.50$ ) or non-parametrically ( $z = 0.49$ ,  $p = 0.63$ ).

**2.2.4 Feature Swap conditions.** As is clear from inspection of figure 4, subjects habituated during the initial familiarization trials later recovered looking time in the Feature Swap test trial, but not in the Feature Swap Control or Position Change Control test conditions. These impressions were verified by the following statistical analyses. During familiarization, subjects looked reliably less in the third familiarization trial than in the first (paired  $t$ -test:  $t_{41} = 7.53$ ,  $p < 0.01$ ). There was a reliable difference between the durations of looking in the three types of test trials ( $F_{2,39} = 6.18$ ,  $p < 0.01$ ), but planned comparisons indicated that the absolute levels of looking in test trials differed



**Figure 4.** Sample stimuli and results from the Feature Swap conditions of experiment 1. The graphs depict mean looking time  $\pm$  standard error for the familiarization and test trials.

only between the Feature Swap and Feature Swap Control test conditions ( $t_{27} = 4.05$ ,  $p < 0.01$ ). The absolute level of looking in test trials was not reliably different between the Position Change Control and Feature Swap Control ( $t_{23} = 1.63$ ,  $p = 0.12$ ) or between the Position Change Control and the Feature Swap condition ( $t_{28} = 1.62$ ,  $p = 0.12$ ). Nevertheless, the three test groups differed in their degree of dishabituation in test trials ( $F_{2,39} = 4.51$ ,  $p = 0.02$ ). Subjects in the Feature Swap condition recovered looking time relative to the final familiarization trial ( $t_{16} = 3.43$ ,  $p < 0.01$ ), but those in the two control conditions did not (Feature Swap Control:  $t_{11} = 1.1$ ,  $p = 0.30$ ; Position Change Control:  $t_{12} = 0.55$ ,  $p = 0.60$ ).

### 2.3 Discussion

The results of this initial experiment demonstrate that rhesus monkeys employ feature binding during passive viewing in a natural environment, without extensive training—or, indeed, any training at all. After only three encounters with an array of geometric objects during the familiarization phase, subjects reliably detected changes made to the colors or shapes of these objects. In addition, the key result of this study was that subjects recovered looking time to a test display which contained exactly the same features as the familiarization display, but combined differently. Because subjects did not recover looking time to the control condition in which the familiarization objects simply swapped relative *locations*, the successful discrimination of the change in feature binding had to have been due to the binding per se, and not simply to the change in location of one of the individual features (either color or shape).

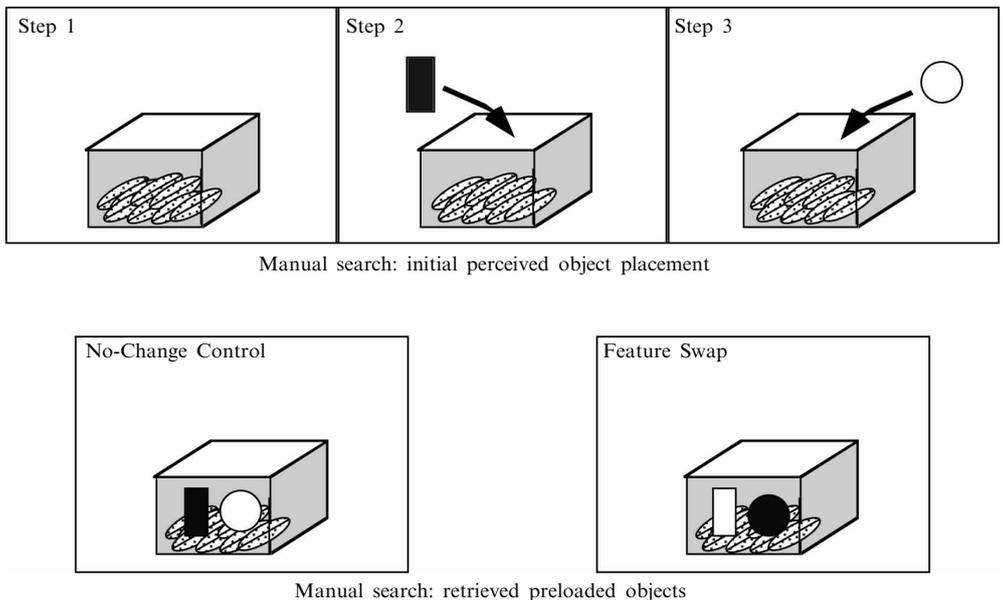
## 3 Experiment 2: Manual search

Previous neurophysiological studies have demonstrated that rhesus monkeys are able to perceive objects as bound collections of features (eg Tanaka 1993; Tanaka et al 1991), and the previous experiment has demonstrated that this process of feature binding also operates spontaneously in a naturalistic setting. None of these studies, however, has addressed the question how (or whether) feature binding is *used* to control real-world behavior. As noted above, there are many examples (especially from the study of human infant cognition) where certain properties are discriminable, but are nevertheless not used in the service of other cognitive operations.

In this experiment we explore the ability of rhesus monkeys to use feature binding in the service of an especially critical perceptual process: tracking objects as the same persisting individuals over time on the basis of their features. Note, in this context, that there are three importantly different ways in which subjects could have interpreted the feature binding change in experiment 1: (i) subjects may have detected that the objects in the test displays were not the same individuals that had been present during familiarization (and may have even wondered where the original objects had gone, assuming they had encountered a total of four objects); (ii) alternatively, subjects may have interpreted the feature swaps as *transformations* in those objects between familiarization and test, and may thus have looked longer because they found the transformations (to those very same objects) to be novel and/or interesting (assuming that they had encountered only two objects); (iii) finally, subjects may have (merely) looked longer at test because they detected the novel feature combination per se, without having any particular percept of inference about whether the objects at test were new or old.

In this experiment we distinguish these options in order to test whether feature binding is not merely computed, but also employed in the service of tracking objects through time and occlusion. We again tested the spontaneous perception of monkeys in a natural environment, but now with a *manual-search* method (Santos et al 2002) which—unlike looking-time—directly taps object individuation (see also Flombaum et al 2004). Monkeys observed two colored pieces of food (eg a white circle and a red

rectangle—both pieces of apple which could be colored with food coloring) placed in a container filled with leaves, after which they were allowed to search the container to retrieve the food items. The key manipulation (between subjects) involved which objects the monkeys retrieved: they either found the same two objects that they had initially seen placed in the container (ie the white circle and the red rectangle), or they found the same objects but with their features swapped (ie a white rectangle and a red circle; see figure 5). What will subjects infer upon encountering the feature-swapped objects in the leaf-filled container? If they detect this swap and interpret it as the retrieval of *two new objects*, then they should continue to search in the container with the goal of also recovering the initial two objects—and they should thus search *longer* than if they retrieved objects that were identical to the initially placed objects. In contrast, if they detect the feature swap but do not use this information in order to individuate and track the persisting identities of the objects, then they may be surprised but should not continue searching after retrieving the (only) two objects they saw placed into the container.



**Figure 5.** Stimulus presentation in experiment 2. A manual search procedure was used in which subjects initially saw two food objects placed in a container filled with leaves. In fact, both objects were surreptitiously removed, and upon searching the container subjects found either (a) two objects which were identical to those seen placed in the container (in the No-Change Control condition); or (b) two objects with their individual shape–color combinations swapped (in the Feature Swap condition).

In this way, we study whether rhesus monkeys will spontaneously *use* feature binding to control a simple foraging behavior in a natural environment.

### 3.1 Method

**3.1.1 Subjects.** We tested 40 rhesus monkeys from the same population as in experiment 1. An additional 35 subjects were thrown out owing to interference from nearby monkeys, inattention during the display, a failure to approach, or experimental error.

**3.1.2 Apparatus.** Subjects were allowed to search for pieces of food inside a small plastic lunch box (15 cm × 25 cm × 15 cm) filled with leaves native to the island. Subjects were allowed to search the box for apple pieces cut into two shapes (a circle, 5 cm × 5 cm; and a rectangle, 3 cm × 6 cm) and two colors (red and white). The red apple pieces

were dyed with edible red food coloring; white pieces were not dyed and contained no peel.

**3.1.3 Procedure.** As in previous manual-search studies with this population (see Santos et al 2002), we chose subjects opportunistically when they were alone and seated. Two experimenters ran each session. The presenter stood approximately 1.5 m in front of the subject and performed all the actions. The second experimenter stood 3 m directly behind the first experimenter and filmed the session.

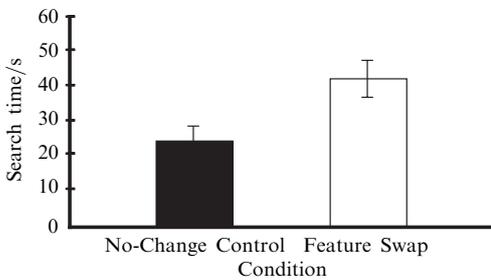
Subjects first witnessed a sequence of events in which apple pieces were placed in a box, as described below. At the beginning of each trial, the presenter crouched to the ground (eye-level with the subject), tilted the box forward to reveal that it contained only leaves, then placed the box on the ground directly in front of him or her. The presenter then reached into a pouch and removed two apple-objects. As in the familiarization used in the feature swap condition of experiment 1, these apple-objects differed along both a shape and color dimension (eg a red rectangle and a white circle). The presenter extended the right arm, displayed an apple-object (eg a red rectangle) to the subject, and then placed it inside the box. While keeping the right hand inside the box, the presenter then extended the left arm, displayed the second apple-object (eg a white circle), and then placed that object (and the hand) inside the box. The order in which the two apple-objects were added to the box was counterbalanced across subjects. The presenter then removed both hands simultaneously (surreptitiously removing the two apple-objects originally placed in the box), stood up, turned away from the subject, and retreated approximately 3 m in the opposite direction towards the second experimenter, allowing the subject to approach and search the box.

In all cases, however, the box was ‘preloaded’ with two distinct apple-objects, which were the objects that the subjects then retrieved at the beginning of their search. Upon searching the box, roughly half of the subjects ( $n = 21$ ) were presented with a No-Change Control event, in which the preloaded (and retrieved) objects were identical to the objects they had seen placed in the container (ie a red rectangle and a white circle). The remaining subjects ( $n = 19$ ), however, were presented with a Feature Swap event, in which the preloaded (and retrieved) objects had the same sets of properties (ie identical colors, shapes, and overall amount of material), but combined differently (ie a red circle and a white rectangle).

The cameraperson, who was blind to the experimental condition, filmed the subject’s approach to the box and subsequent searching. The cameraperson continued filming until the subject had voluntarily finished searching. The cameraperson decided that the subject finished searching when it had moved more than 10 m from the box and/or had stopped reaching into the container for a continuous 60 s. The videotapes were later recorded onto an iBook laptop computer and were analyzed with Adobe Premiere software. A single experimenter who was blind to the test condition coded subjects’ search time during each 33 ms frame of the period between the retrieval of the second preloaded object and the end of the trial. Searching was defined as a period in which the subject had its hand inside the box and/or its head oriented toward the inside of the box. One quarter of the trials in each condition were independently scored by a second coder; the two coders were highly reliable ( $r = 0.99$ ), and the first coder’s judgments were used in case of a disagreement.

### 3.2 Results and discussion

As depicted in figure 6, subjects in the Feature Swap condition searched longer (42.17 s) than subjects in the No-Change Control condition (24.56 s), a difference that was reliable both parametrically ( $t_{38} = 2.10$ ,  $p = 0.03$ ) and non-parametrically ( $z = 2.28$ ,  $p = 0.02$ ).



**Figure 6.** Search time with standard errors in experiment 2, measured from the moment that the second food object was retrieved.

Subjects in this experiment were confronted with a container in which they had seen two food objects placed, and from which they later retrieved two food objects. When these two retrieved food objects had the identical features to those of the objects initially placed in the box, but in a novel combination, subjects tended to keep searching—presumably for the additional ‘missing’ original food objects—more than 70% longer than when the initially placed pair of objects and the retrieved pair of objects were identical. As in experiment 1, these results demonstrate that rhesus monkeys are able to spontaneously discriminate not only novel features, but also novel feature *combinations*. In addition, the results of this experiment go further, demonstrating that monkeys will also *use* representations of bound features in the service of other cognitive processes, such as object individuation.

#### 4 General discussion

The results of the two experiments presented here demonstrate that at least one non-human primate—the rhesus macaque—is able to spontaneously bind visual features into coherent object representations. This result by itself is perhaps unremarkable, given that in previous neurophysiological work with monkeys feature binding was also found after considerable training. Moreover, it is perhaps unsurprising that *any* animal with a developed visual system should experience bound features given that the world itself contains discrete objects.

In experiment 2, however, rhesus macaques not only detected the presence of a feature swap between two objects, but also *used* this information to infer the presence of new objects in their environment: the monkeys in this experiment apparently inferred on the basis of the change in feature binding that the objects they retrieved from the container were distinct from those they had initially seen placed into the container. In contrast, if they had simply interpreted this event as involving a *transformation* of the initial objects—as may have been the case in experiment 1—then this could have led to surprise, but should not have led to increased searching.

This demonstration that feature binding is used in the service of other perceptual and cognitive processes is more noteworthy than the simple (though novel) demonstration of spontaneous feature binding. This point is perhaps made most clearly in the aforementioned studies of infants’ representations of persisting objects. Although very young infants perceive features of an object well enough to visually segment simple displays using color and shape information (Needham 1998; Wilcox 1999), and even seem to spontaneously bind this information to discrete object representation (Bushnell and Roder 1985; Kaldy and Leslie 2003; Slater et al 1991; Taga et al 2002), their ability to use these properties to track and identify objects as either new or old does not develop until months later (Tremoulet et al 2000; Xu and Carey 1996). Similarly, it could have been the case that rhesus monkeys would perceive features and their combinations, but would not use this information to control individuation—especially in complex natural environments. A similar phenomenon seems to hold for the perception of mechanical properties, such as solidity, by rhesus monkeys. Rhesus monkeys

readily detect perceptual violations of solidity (eg one object moving through a solid barrier; see Santos and Hauser 2002), but fail to employ this mechanical constraint when predicting the motion of objects in foraging situations (Hauser 2001; Santos 2004).

This demonstration of the role of feature binding in object individuation in rhesus monkeys is informative for several related literatures. Our results add to previous neurophysiological research on monkeys by showing that feature binding occurs spontaneously, without extensive training. Our results add to previous work on human adults and young infants by demonstrating how feature binding may contribute in natural environments to other types of natural behaviors such as simple foraging. Indeed, it would be useful to adapt the manual-search method employed in experiment 2 for an analogous study with human infants. Many previous infant studies have explored the basic *discriminability* of feature swaps (eg Bushnell and Roder 1985; Kaldy and Leslie 2003; Slater et al 1991; Taga et al 2002), and many other studies have explored the nature of infants' object individuation (see Carey and Xu 2001), but, to our knowledge, no previous studies have explored whether human infants will use feature binding in the service of object individuation.

#### 4.1 Conclusion

Our visual experience—and the world itself—consists of discrete objects, each of which has its own features: we may see a room containing a brown chair, a white table, and a silver computer. This needn't be the case. Indeed, the binding between object shapes and colors may become fluid in certain neuropsychological disorders, such that patients could visually experience a chair which is initially brown, then white, then silver (Friedman-Hill et al 1995; Robertson et al 1997). Even in normal individuals, certain conditions can frustrate this binding, so that subjects might (mis-)perceive a white chair, a silver table, and a brown computer. Moreover, perceiving any binding at all is an accomplishment, given that features are initially processed in separate functional and anatomical streams: without a process of feature binding, percepts might consist of an unstructured set of unbounded shapes (a chair, a table, and a computer) and amorphous color patches (brown, silver, and white). This is perhaps characteristic of very young infants' perception of the world: though young infants (around 1 month of age) can discriminate feature combinations, later stages of development (even a month later) seem to disrupt this ability under the same circumstances (Taga et al 2002).

For all of these reasons, feature binding—while a pervasive element of our visual experience—is (like most visual processes) an impressive accomplishment. The present results demonstrate that rhesus monkeys are also accomplished in this respect: they not only bind features into coherent object representations, but they also spontaneously use feature binding to individuate objects over time.

**Acknowledgments.** For helpful conversation and/or comments on earlier drafts we thank Marvin Chun and Jon Flombaum. We also wish to thank Elyssa Berg, Madeline Kerner, Webb Phillips, Wendy Reilly, and April Ruiz for their help running these studies, and Melissa Gerald for her help in securing the Cayo Santiago field station. BJS was supported by NSF #BCS-0132444. The Cayo Santiago field station was supported by the NIH (NCRR-5P40RR03640).

#### References

- Alvarez G A, Cavanagh P, 2004 “The capacity of visual short-term memory is set both by visual information load and by number of objects” *Psychological Science* **15** 106–111
- Bushnell E W, Roder G J, 1985 “Recognition of color/form compounds by 4-month-old infants” *Infant Behavior and Development* **8** 255–268
- Carey S, Xu F, 2001 “Infant knowledge of objects: Beyond object files and object tracking” *Cognition* **80** 179–213
- Chen L, Zhou W, 1997 “Holes in illusory conjunctions” *Psychonomic Bulletin & Review* **4** 507–511
- Cohen A, Ivry R B, 1989 “Illusory conjunctions inside and outside the focus of attention” *Journal of Experimental Psychology: Human Perception & Performance* **15** 650–662

- Desimone R, Albright T D, Gross C G, Bruce C, 1984 "Stimulus-selective properties of inferior temporal neurons in the macaque" *Journal of Neuroscience* **4** 2051–2062
- Desimone R, Schein S, 1987 "Visual properties of neurons in area V4 of the macaque: Sensitivity to stimulus form" *Journal of Neurophysiology* **57** 835–868
- Desimone R, Schein S J, Moran J, Ungerleider L G, 1985 "Contour, color and shape analysis beyond the striate cortex" *Vision Research* **25** 441–452
- DeYoe E A, Van Essen D C, 1988 "Concurrent processing streams in monkey visual cortex" *Trends in Neurosciences* **115** 219–226
- Engel A K, Singer W, 2001 "Temporal binding and the neural correlates of sensory awareness" *Trends in Cognitive Sciences* **5** 16–25
- Flombaum J I, Junge J A, Hauser M D, 2005 "Rhesus monkeys (*Macaca mulatta*) spontaneously compute additional operations over large numbers" *Cognition* **97** 315–325
- Flombaum J I, Kunder S M, Santos L R, Scholl B J, 2004 "Dynamic object individuation in rhesus macaques: A study of the tunnel effect" *Psychological Science* **15** 795–800
- Friedman-Hill S R, Robertson L C, Treisman A, 1995 "Parietal contributions to visual feature binding: evidence from a patient with bilateral lesions" *Science* **269** 853–855
- Gattass R, Sousa A P, Gross C G, 1988 "Visuotopic organization and extent of V3 and V4 of the macaque" *Journal of Neuroscience* **8** 1831–1845
- Gawne T J, Kjaer T W, Richmond B J, 1996 "Latency: another potential code for feature binding in striate cortex" *Journal of Neurophysiology* **76** 1356–1360
- Hauser M D, 2001 "Searching for food in the wild: A nonhuman primate's expectations about invisible displacement" *Developmental Science* **4** 84–93
- Hauser M D, MacNeilage P, Ware M, 1996 "Numerical representations in primates" *Proceedings of the National Academy of Sciences of the USA* **93** 1514–1517
- 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
- Kahneman D, Henik A, 1981 "Perceptual organization and attention", in *Perceptual Organization* Eds M Kubovy, J Pomerantz (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 181–211
- Kaldy Z, Blaser E, Leslie A, 2004 "Iso-salient color and luminance information in object identification" poster presented at the *International Conference on Infant Studies*, 5/5/05–5/8/05, Chicago, IL
- Kaldy Z, Leslie A M, 2003 "Identification of objects in 9-month-old infants: integrating 'what' and 'where' information" *Developmental Science* **6** 360–373
- Livingstone M S, Hubel D H, 1984 "Specificity of intrinsic connections in primate primary visual cortex" *Journal of Neuroscience* **4** 2830–2835
- Livingstone M S, Hubel D H, 1987 "Psychophysical evidence for separate channels for the perception of form, color, movement, and depth" *Journal of Neuroscience* **7** 3416–3468
- Livingstone M S, Hubel D H, 1988 "Segregation of form, color, movement, and depth: Anatomy, physiology, and perception" *Science* **240** 740–749
- Luck S J, Chelazzi L, Hillyard S A, Desimone R, 1997 "Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex" *Journal of Neurophysiology* **77** 24–42
- Luck S, Vogel E, 1997 "The capacity of visual working memory for features and conjunctions" *Nature* **390** 279–281
- Moran J, Desimone R, 1985 "Selective attention gates visual processing in the extrastriate cortex" *Science* **229** 782–784
- Müller H J, Elliott M A, Herrmann C S, Mecklinger A (Eds), 2001 *Neural Binding of Space and Time: Spatial and Temporal Mechanisms of Feature–Object Binding* (Hove, East Sussex: Psychology Press)
- Munakata Y, Santos L, O'Reilly R, Hauser M D, Spelke E S, 2001 "Visual representation in the wild: how rhesus monkeys parse objects" *Journal of Cognitive Neuroscience* **13** 44–58
- Needham A, 1998 "Infants' use of featural information in the segregation of stationary objects" *Infant Behavior and Development* **21** 47–76
- O'Craven K, Downing P, Kanwisher N, 1999 "fMRI evidence for objects as the units of attentional selection" *Nature* **401** 584–587
- Prinzmetal W, Henderson D, Ivry R, 1995 "Loosening the constraints on illusory conjunctions: The role of exposure duration and attention" *Journal of Experimental Psychology: Human Perception and Performance* **21** 1362–1375
- Rawlins R, Kessler M, 1987 *The Cayo Santiago Macaques* (New York: State University of New York Press)

- 
- Reynolds J, Chelazzi L, Desimone R, 1999 "Competitive mechanisms subserve attention in macaque areas V2 and V4" *Journal of Neuroscience* **19** 1736–1753
- Riesenhuber M, Poggio T, 1999 "Are cortical models really bound by the 'binding problem'?" *Neuron* **24** 87–93
- Robertson L, Treisman A, Friedman-Hill S, Grabowecky M, 1997 "The interaction of spatial and object pathways: Evidence from Balint's syndrome" *Journal of Cognitive Neuroscience* **9** 295–317
- Santos L R, 2004 "'Core Knowledge': a dissociation between spatiotemporal knowledge and contact-mechanics in a non-human primate?" *Developmental Science* **7** 164–174
- Santos L R, Hauser M D, 2002 "A nonhuman primate's understanding of solidity: Dissociations between seeing and acting" *Developmental Science* **5** F1–F7
- Santos L R, Miller C T, Hauser M D, 2003 "Representing tools: how two non-human primate species distinguish between the functionally relevant and irrelevant features of a tool" *Animal Cognition* **6** 269–281
- Santos L R, Sulikowski G M, Spaepen G M, Hauser M D, 2002 "Object individuation using property/kind information in rhesus macaques (*Macaca mulatta*)" *Cognition* **83** 241–264
- Singer W, Gray C M, 1995 "Visual feature integration and the temporal correlation hypothesis" *Annual Review of Neuroscience* **18** 555–586
- Slater A, Mattock A, Brown E, Burnham D, Young A, 1991 "Visual processing of stimulus compounds in newborn infants" *Perception* **20** 29–33
- Taga G, Ikejiri T, Tachibana T, Shimojo S, Soeda A, Takeuchi K, Konishi Y, 2002 "Visual feature binding in early infancy" *Perception* **31** 273–286
- Tanaka K, 1993 "Neuronal mechanisms of object recognition" *Science* **262** 685–688
- Tanaka K, 1996 "Inferotemporal cortex and object vision" *Annual Review of Neuroscience* **19** 109–139
- Tanaka K, Saito H, Fukada Y, Moriya M, 1991 "Coding of visual images of objects in the inferotemporal cortex of the Macaque monkey" *Journal of Neurophysiology* **66** 170–189
- Thiele A, Stoner G, 2003 "Neuronal synchrony does not correlate with motion coherence in cortical MT" *Nature* **421** 366–370
- Treisman A M, 1988 "Features and objects: The Fourteenth Bartlett Memorial Lecture" *Quarterly Journal of Experimental Psychology* **40** 201–237
- Treisman A M, 1999 "Solutions to the binding problem: Progress through controversy" *Neuron* **24** 105–110
- Treisman A M, Schmidt H, 1982 "Illusory conjunctions in the perception of objects" *Cognitive Psychology* **14** 107–141
- Tremoulet P, Leslie A M, Hall G, 2000 "Infant attention to the shape and color of objects: Individuation and identification" *Cognitive Development* **15** 499–522
- Van de Walle G, Carey S, Prevor M, 2000 "Bases for object individuation in infancy: Evidence from manual search" *Journal of Cognition and Development* **1** 249–280
- Wilcox T, 1999 "Object individuation: infants' use of shape, size, pattern, and color" *Cognition* **72** 125–166
- Wu D-A, Kanai R, Shimojo S, 2004 "Vision: Steady-state misbinding of colour and motion" *Nature* **429** 262
- Xu F, Carey S, 1996 "Infants' metaphysics: the case of numerical identity" *Cognitive Psychology* **30** 111–153



ISSN 0301-0066 (print)

ISSN 1468-4233 (electronic)

# PERCEPTION

VOLUME 35 2006

[www.perceptionweb.com](http://www.perceptionweb.com)

**Conditions of use.** This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.