



The “what” and “where” of object representations in infancy

Denis Mareschal*, Mark H. Johnson

Centre for Brain and Cognitive Development, School of Psychology, Birkbeck College, University of London, Malet Street, London, WC1E 7HX, UK

Received 6 February 2002; received in revised form 6 September 2002; accepted 3 January 2003

Abstract

Four-month-olds' memory for surface feature and location information was tested following brief occlusions. When the target objects were images of female faces or monochromatic asterisks infants showed increased looking times following a change in identity or color but not following a change in location or combinations of feature and location information. When the target objects were images of manipulable toys, the infants showed increased looking times following a change in location but not identity or the binding of location and identity information. This evidence is consistent with the idea that young infants are unable to maintain the information processed separately in both the dorsal and ventral visual streams during occlusions. Our results suggest that it is the target's affordance for action that determines whether the dorsal or ventral information is selectively maintained during occlusion.

© 2003 Elsevier Science B.V. All rights reserved.

Keywords: Infancy; Dorsal/ventral streams; Occlusion

1. Introduction

Preferential looking studies have found that infants will show increased looking time at occlusion events in which some property of a temporarily hidden object is altered.¹ Examples of properties tested in this way have included the violation of object location,

* Corresponding author.

E-mail address: d.mareschal@bbk.ac.uk (D. Mareschal).

¹ Increased looking time is often interpreted as evidence of surprise due to a violation of the infant's expectation of how objects in the world behave. In the work below, we make only the minimal assumption that infants are endowed with a memory that maintains sufficient information to drive a novelty preference response to some aspect of the test events for at least the occlusion duration in our studies.

size, shape, volume, number, and trajectory (e.g. Baillargeon, 1993; Spelke, Katz, Purcell, Ehrlich, & Breilinger, 1994; Wynn, 1992).

Studies that rely on similar occlusion paradigms to assess the basis of object individuation have provided a mixed picture of what features infants can use to individuate hidden objects. At times, infants appear not to use object features to individuate objects (Xu & Carey, 1996, 2000; Xu, Carey, & Welch, 1999). However, in other studies they do appear to process feature information, and can use this information to individuate objects (Tremoulet, Leslie, & Hall, 2000; Wilcox, 1999; Wilcox & Baillargeon, 1998; Wilcox & Schweinle, 2002). Indeed, Wilcox (1999) reports that infants can use shape at 7 months, texture at 11 months, and color at 12 months to individuate objects. Taken together, this evidence has led to the suggestion that early infant object representations are mainly spatial and temporal in nature, or at the very least, that infants preferentially encode spatial and temporal information over and above surface feature information (Leslie, Xu, Tremoulet, & Scholl, 1998; Wilcox, 1999; Wilcox & Schweinle, 2002; Xu, 1999; Xu & Carey, 2000).

However, although color, size, shape, texture, velocity and trajectory are all traits of an individual object, not all traits are processed in the same way in the brain. There is a natural divide that follows from a consideration of the neurophysiology of visual object processing. Substantial evidence has been mustered to suggest that there are two routes for visual object processing (Goodale, Milner, Jakobson, & Carey, 1991; Milner & Goodale, 1995; Ungerleider & Mishkin, 1982; van Essen, Anderson, & Felleman, 1992). Although the exact function of these streams is still debated (Johnson, Mareschal, & Csibra, 2001), it is generally agreed that the dorsal route largely processes spatial and temporal object information involved in guiding action (e.g. location, motion, size, crude shape for grasping, but *NOT* color or face information²) (Jeannerod, 1988; Newsome, Wurtz, & Komatsu, 1988; Rizzolatti, Riggio, & Sheliga, 1994; Taira, Mine, Georgopoulos, Mutara, & Sakata, 1990). In contrast, the ventral route processes information relevant for the identification of objects (e.g. color, face information, and *ALSO* size, shape, but *NOT* location; Gross, 1992; Komatsu, Ideura, Kaji, & Yamane, 1992; Perret, Mistlin, & Chitty, 1987).

Mareschal, Plunkett, and Harris (1999) argued that young infants are poor at integrating information processed separately in the dorsal and ventral cortical pathways. This weakness becomes apparent when infants are required to respond to occluded objects. When objects are occluded, infants must not only maintain representations of the object's surface features and location, but they must also maintain the binding between the two representations in different cortical streams. This applies to reaching responses, but also to looking responses traditionally associated with earlier competence. Mareschal et al. argued that infants appeared to perform better in visual preference tasks (in which looking time is the measured variable) than manual retrieval tasks because existing visual preference tasks did not require infants to integrate information across the dorsal and ventral cortical routes, whereas most manual retrieval tasks did. This is because, in manual retrieval tasks, the infant needs to recognize the object as a desired object in order to engage in a reach response (Rao, Rainer, & Miller, 1997). An implication of this argument is that infants will

² Indeed, Milner and Goodale (1995, p. 134) use the fact that DF had preserved color processing to argue that she has a partially functional ventral stream precisely *because* the dorsal stream does not process color.

show a developmental delay, similar to that observed in manual retrieval tasks, on looking time tasks that *do* require the integration of dorsal and ventral object information. Tasks that require infants to detect a violation of the binding of dorsal and ventral object information are one such example.

In this paper, we explore further the idea that young infants are poor at coordinating and integrating dorsal and ventral object representations. Specifically, we predict that 4-month-olds will provide evidence of both functional dorsal and ventral route processing following brief occlusions, but no evidence of integration of information across the two routes. We focus on 4-month-olds because they are the youngest age group at which evidence of a response to the violation of some property of a hidden object has been consistently provided (e.g. Baillargeon, 1993; Spelke et al., 1994).

To assess these predictions, we use a brief occlusion paradigm in which key traits of a hidden object are changed during the occlusion period. We are careful to choose object traits that show a clear dissociation between ventral route (e.g. color, faces) and dorsal route (e.g. location) processing. Shape and size are not good discriminators because they are thought to be processed to some extent by both pathways. In other words, shape and location binding can be occurring within the dorsal route only, and may not require the concurrent access to the ventral stream.³ Indeed, recent evidence (Tremoulet et al., 2000) suggests that infants can form and maintain feature bindings that do not cross the cortical-stream divide (e.g. shape and location).

In addition, we selected target stimuli that would differentially activate dorsal and ventral routes. Gibson (1979) argued that the central function of vision is to provide information about the possibilities of actions – something which he referred to as the “affordance” of an object. Viewing an object that supports action is sufficient to prime motor responses appropriate for that object (Craighero, Fadiga, Rizzolatti, & Umiltà, 1999; Tucker & Ellis, 1998). PET studies have provided converging evidence that the simple perception of objects automatically affords actions that can be made towards them (e.g. Grèze & Decety, 2002, found that passively viewing images of household objects that afford action activated the dorsal pathways in adults). The translation of visual information into an appropriate motor code happens in the dorsal system (see also Milner & Goodale, 1995). Tucker and Ellis (2001) suggest that, at some basic level, this process is automatic and independent of any high-level goals. This is because goals are usually specified at a relatively high level (e.g. make a cup of coffee). How the details of the high-level goals are formed depends on an ability to pick up information about the possible actions in a scene, that may or may not contribute to the desired end state. Under conditions where goals are under-specified, or loosely specified, there must be a process that enables the most relevant vision-to-motor transformation to take priority. If the perception of action affordances is automatic and independent of high-level goals then it seems reasonable to ask whether such processes may be in operation early in infancy.

In sum, the experiment described below was designed to explore (1) whether young

³ There are cells that respond to object shape in both the parietal and inferior temporal lobes (see Jeannerod, 1988; Milner & Goodale, 1995). However, whether this response is due to cross-talk between the pathways is not important for our argument. What is important is that shape does not clearly differentiate between dorsal and ventral processing.

infants could encode and maintain object feature *and* location information over brief delays, and (2) whether the affordance of the target interacted with their response.

2. Methods

2.1. Participants

Eighty infants (35 girls and 45 boys) took part in this study. All infants were tested within 8 days on either side of their 4-month birthdays. The median age of the sample was 4 months and 2 days, and all infants had a gestation period of between 38 and 42 weeks. The participants were predominantly from middle class Caucasian backgrounds. A further 20 infants were tested but were not included in the analysis because of fussiness (two), a failure to look a minimum of 1 s in each of the four test conditions (twelve), experimenter error/equipment failure (five), and mother intervention (one).

2.2. Procedure

Infants were tested at the Centre for Brain and Cognitive Development. Parents were led into a reception/play room in which the study was explained to them and they signed consent forms. They were then led into a special purpose testing room painted gray, and illuminated from above. The infants were sat in a car seat facing a 21 inch computer monitor located 60 cm in front of them. The monitor was encased in a black curtain. A dull gray partition was positioned to isolate the child from the parent and experimenter, and to hide the experimental control center from the infant's view. All events were computer controlled and presented on this monitor. This ensured that stimulus characteristics such as target velocities and occlusion rates were constant across all infants and in all experimental conditions.

The familiarization phase consisted of five repetitions of the same events (Fig. 1) sequenced automatically by the computer, in which a digital target image moved out from behind an occluder, moved towards the edge of the monitor, remained stationary for 5 s, and then moved back behind the occluder. As soon as this first target had disappeared, a second target moved out from behind a second occluder, moved towards the opposite edge of the monitor, remained stationary for 5 s, and then returned behind the occluder. The appearance and disappearance of both targets constituted a complete familiarization trial (Fig. 1, left panel).

During the familiarization phase, no object ever passed in the gap between the occluders. Moreover, the disappearance of one target and the appearance of another target were temporally contiguous. This provided an additional cue to rule out the possibility that a single target was traversing the distance behind the occluders. During motion, all targets moved at a constant 9 cm/s. In addition, black shading along the outer border of the occluders provided some further evidence of depth in the image.

Following the familiarization trials, the test trial *per se* began. Each infant saw four different test trials. Presentation order was determined by a repeated Latin-Square design. A test trial began when the infant was fixating a green flashing centering cue between the two yellow occluders (the flashing cue disappeared as soon as the trial began). Between

trials, a noise was made to help re-orient the infant to the center of the monitor. Each test trial began with one familiarization trial (i.e. one target moving in and out from behind each occluder). At the end of familiarization trial (i.e. once the second target had returned behind its original occluder) there was a 5 s pause. After the pause, both yellow occluders were translated smoothly upwards (by scrolling the occluder images) towards the top of the monitor at 5 cm/s, and revealed what lay behind them (Fig. 1, right panel).

In the *Baseline* test trial, the occluders moved up to reveal the two targets from the familiarization phase, each behind the occluder from which it had appeared during the familiarization phase. Looking times at this event provide a baseline for how engaging the event is when there has been no violation of feature or location information. All looking times during other test trials were compared to this baseline value.

In the *Surface Feature* (SF) test trial, the occluders moved up to reveal one target behind

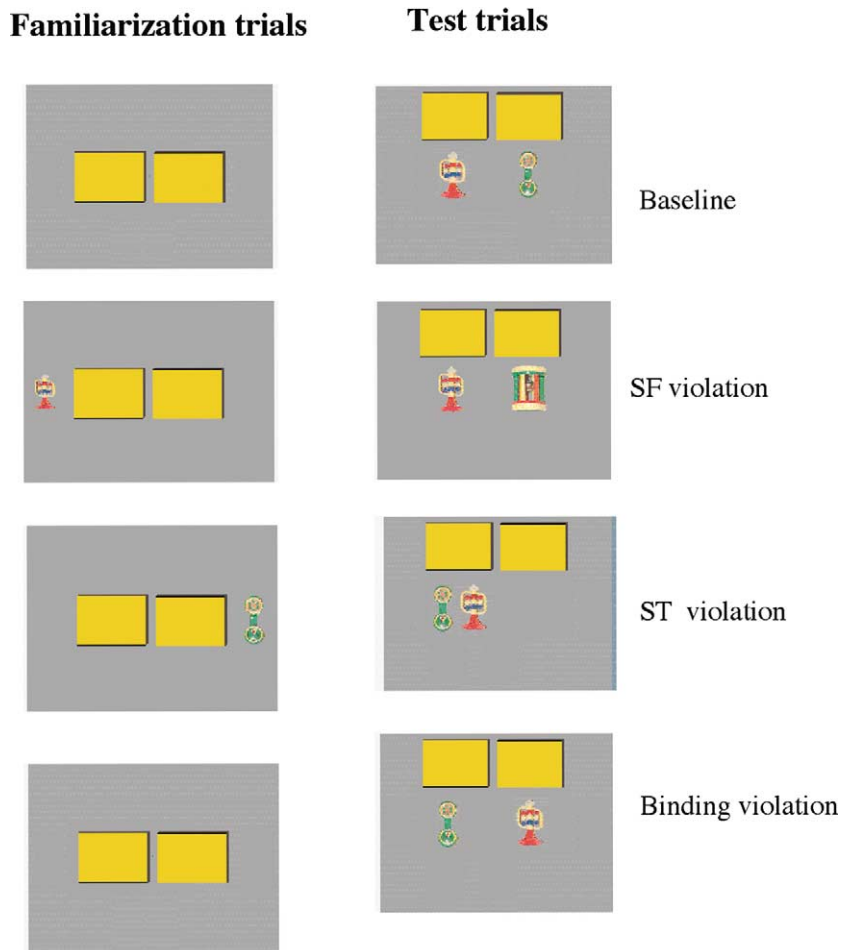


Fig. 1. Familiarization (left) and test events (right).

each occluder. However, one of the two targets was novel while the other was a familiar target behind the expected occluder. This event contains a violation of feature information (novel target), but not location information (there is one target behind each occluder). Increased looking times to this event (over and above looking times in the Baseline condition) would suggest that the infants encoded the features of the targets, remembered the features of the targets during the occlusion period, and responded to the novelty of the new target.

In the *Spatial-Temporal* (ST) test trial, the occluders moved up to reveal no target behind one occluder and both targets behind the other occluder. Moreover, the final locations of the targets were such that they would have had to cross over one another. This event contains a violation of spatial-temporal (location) information but no violation of feature information (same targets). Increased looking to this event (over and above looking times in the Baseline condition) would suggest that the infants encoded the location of targets, remembered the location of the targets during the occlusion period, and responded to the novelty of the target locations.

In the *Binding test trial* the occluders moved up to reveal one of the targets from the familiarization phase behind each occluder. However, the targets had been switched around such that each target was behind the occluder from which it had never appeared. This condition contains a violation of the binding (conjunction) between surface-feature and spatial-temporal (location) information. The separate components in this display (surface features and location) are identical to those in the Baseline condition but the feature-location compounds have changed. Increased looking times to this event (over and above looking times in the baseline test trial) would suggest that the infants encoded the locations of targets, processed the features of targets, remembered the location and features of the targets, remembered which target was at which location during the occlusion period, and responded to the novelty of the feature-location compound.

In all test trials, the occluding screens remained raised and stationary for a maximum of 15 s after which they descended to their original positions, thereby re-occluding the targets. The total possible target viewing time (including periods of partial occlusion) was 21 s. Test trials were terminated if infants looked away from the event for more than 3 s consecutively. To terminate a trial, the experimenter pushed a computer key that lowered the occluding screens as though the trial was terminating naturally. The left–right location of the novel target (in the SF test trial) and the empty occluder (in the ST test trial) were each counter-balanced across infants, and independently of one another.

Infants were allotted to one of four experimental conditions. In all conditions the targets were two-dimensional digital images. In the *Faces* condition, the targets were three disembodied female faces randomly selected from a set of seven faces (Fig. 2a) drawn from the McMaster data-base. In the *Asterisks* condition, the targets were cartoon monochromatic asterisks. For each participant, three different asterisks were selected randomly (two familiarization asterisks and one additional asterisk for the SF test trial) from a pool of six monochromatic asterisks (Fig. 2b). These were generated in Adobe PhotoShop and had the following RGB co-ordinates (Blue: 0, 51, 204; Red: 204, 0, 51; Purple: 204, 0, 204; Green: 0, 204, 51; Brown: 153, 102, 0; Orange: 255, 102, 0).

Both face and color information are known to be processed by the ventral stream (Milner & Goodale, 1995). Moreover, two-dimensional faces and flat cartoons do not

naturally afford action so we expect these stimuli to be diagnostic of ventral processing. The two remaining experimental conditions involve object images that we assumed could potentially afford actions.

In the *Toys with manipulation* condition, we used digital images of common graspable infant toys. Pictures of these toys were generated using a Kodak Digital Science DC50 zoom camera. Three toy images were drawn randomly (two familiarization toys and one additional toy for the SF test trial) from a pool of seven images (Fig. 2c). We reasoned that if the dorsal/ventral processing divide reflects a recognition vs. action dissociation, then a target that affords an obvious action may provide a bias towards dorsal processing.

To increase the likelihood that infants would respond to these images as objects that afford actions, we provided them with prior experience of the actual toys in action. Before going into the testing suite, the experimenter presented and demonstrated the use of the actual test toys to each of the infants. The experimenter sat off to one side of the baby, and presented the toy approximately 15 cm in front of the infant. The infant's attention was drawn to the toy, and the toy was shaken, rattled, or turned as appropriate. Each toy was presented for a minimum of 10 s and a maximum of 30 s. Once the babies had seen the three toys once, all three toys were presented a second time.⁴ Presentation order was random and unrelated to the subsequent presentation order in the study. The purpose of this manipulation was to encourage the infant to form a representation of the object image (in the subsequent test phase) as that of a real object that could support action in a three-dimensional world.

Finally, in the *Toys without manipulation* condition, infants were presented with three randomly selected toy images selected from the same pool as above, but, as in the first two experimental conditions, the infants were provided with no prior experience of the objects. The results of this condition will enable an assessment of the extent to which behavioral differences observed in the Toy with manipulation condition reflect the impact of the prior familiarization rather than the content of the image itself.

The monitor viewing area was 38.5×29.5 cm and of a dull gray color (RGB: 153, 153, 153). The yellow occluders (RGB: 255, 204, 0) were 11×7 cm with a 1 cm gap between them.⁵ The target images varied in size slightly depending on the object, but were constrained to be within a 4×6 cm rectangle. Thus, any two objects could legitimately fit behind a single occluder. During the familiarization phase, the targets moved to a position where their closest edge was 3 cm away from the closest edge of the occluder.

2.3. Coding

Infant looking time to the display was coded off-line. During the familiarization trials the scoring interval started as soon as the first target began to appear from behind an

⁴ Because reaching does not normally begin until 4.5 months of age (Gordan, 2001), the majority of infants did not touch the toys but simply watched them being manipulated.

⁵ Due to constraints imposed by the dimensions of the computer monitor, these dimensions differed somewhat from those used in previously published individuation studies. For comparisons, Xu and Carey (1996) used two 11×25 cm occluders separated by a 12 cm gap for a viewing distance of 66 cm, whereas Wilcox and Schweinle (2002) used two 11.5×36 cm occluders separated by a 4 cm gap for a viewing distance of 80 cm. Of course, our paper does not directly address the issue of individuation.

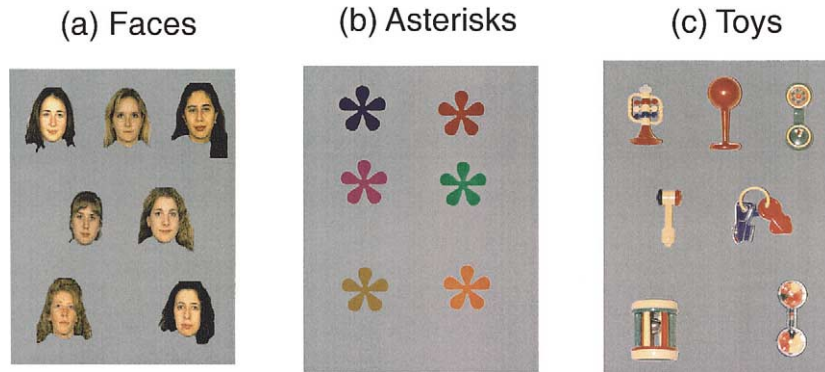


Fig. 2. Target stimuli in the (a) Faces, (b) Asterisks, and (c) Toys experimental conditions.

occluder and finished as soon as the second target had completely disappeared behind its occluder. During the test trials, the scoring interval began as soon as the occluders began to reveal partially occluded objects and ended as soon as the descending occluders had completely occluded the targets. A second independent coder scored 25% of the trials. The Pearson correlation between the two coders' looking times was 0.95 and is typical of such measures in infant looking time studies (Haaf, Brewster, de Saint-Victor, & Smith, 1989).

3. Results

3.1. Familiarization trials

Table 1 shows the mean total looking time during each of the final four familiarization trials in each of the four experimental conditions.⁶ Total looking times were entered into an analysis of variance (ANOVA) with two factors: Trial (with four levels) as a within-subject factor and Condition (with four levels) as a between-subjects factor. The ANOVA revealed a significant main effect of Condition ($F(3, 76) = 3.22, P < 0.03$) and a significant main effect of Trial ($F(3, 228) = 6.233, P < 0.0005$) but no significant interaction of Trial \times Conditions ($F(9, 228) = 0.918, P > 0.5$). The main effect of Trial could be explained by the presence of a significant decreasing linear trend ($F(1, 228) = 11.33, P < 0.001$) and a significant negative quadratic trend ($F(1, 228) = 5.88, P < 0.02$). Thus, there is a decrease in total looking time as familiarization trials unfold. The Condition main effect could be explained by the fact that looking times in the Face condition were significantly greater than those in either the Toys with manipulation condition ($t(19) = 2.15, P < 0.04$) or the Toys without manipulation conditions ($t(19) = 2.94, P < 0.005$). There was also a non-significant trend towards greater looking in the Aster-

⁶ For 29 of the 80 participants, the initial portion of the first familiarization trial was not recorded because of a technical fault. Because we could not reconstruct the full looking times to the first familiarization trial for all participants, we have decided to exclude this trial from any further analyses.

Table 1
Mean total looking times (and standard deviations) in the last four familiarization trials

Condition	Trial 2	Trial 3	Trial 4	Trial 5
Faces	5.72 (1.03)	5.74 (0.83)	5.63 (1.23)	5.42 (1.24)
Colors	5.24 (1.63)	5.67 (0.90)	5.55 (1.52)	4.24 (1.96)
Toys with manipulation	5.00 (1.63)	4.62 (2.36)	5.07 (2.071)	4.39 (2.19)
Toys without manipulations	4.98 (1.70)	4.80 (1.83)	4.46 (1.96)	3.67 (2.53)

isks condition than the Toys without manipulation condition ($t(19) = 1.78$, $P < 0.08$) with no other looking times differences between conditions approaching significance (all $t < 1.15$, $P > 0.27$).

These findings suggest that there is a general decrease in looking across familiarization trials that is present in all four conditions, suggesting that infants are encoding some aspect of the display in all conditions. While the pattern of decrement does not change between conditions, there was slightly greater total looking in the Faces condition than in the two Toys conditions, but no significant differences in total looking time between the Faces and Asterisks conditions or between the Asterisks condition and the two Toys conditions. The slightly greater looking times in the Faces conditions can be attributed to the well documented preference in young infants to look at Faces as opposed to other stimuli (Johnson & Morton, 1991).

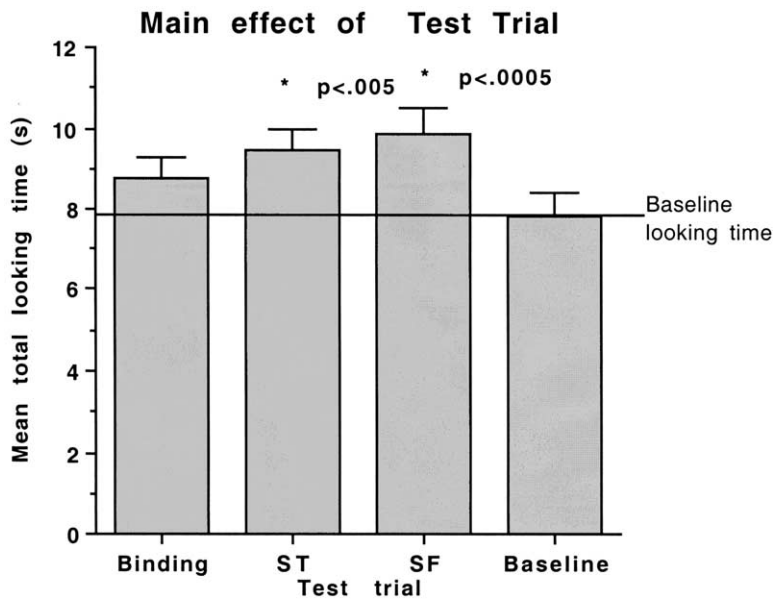


Fig. 3. Mean total looking times in different test trials collapsed across experimental conditions ($n = 80$). Standard errors are also shown.

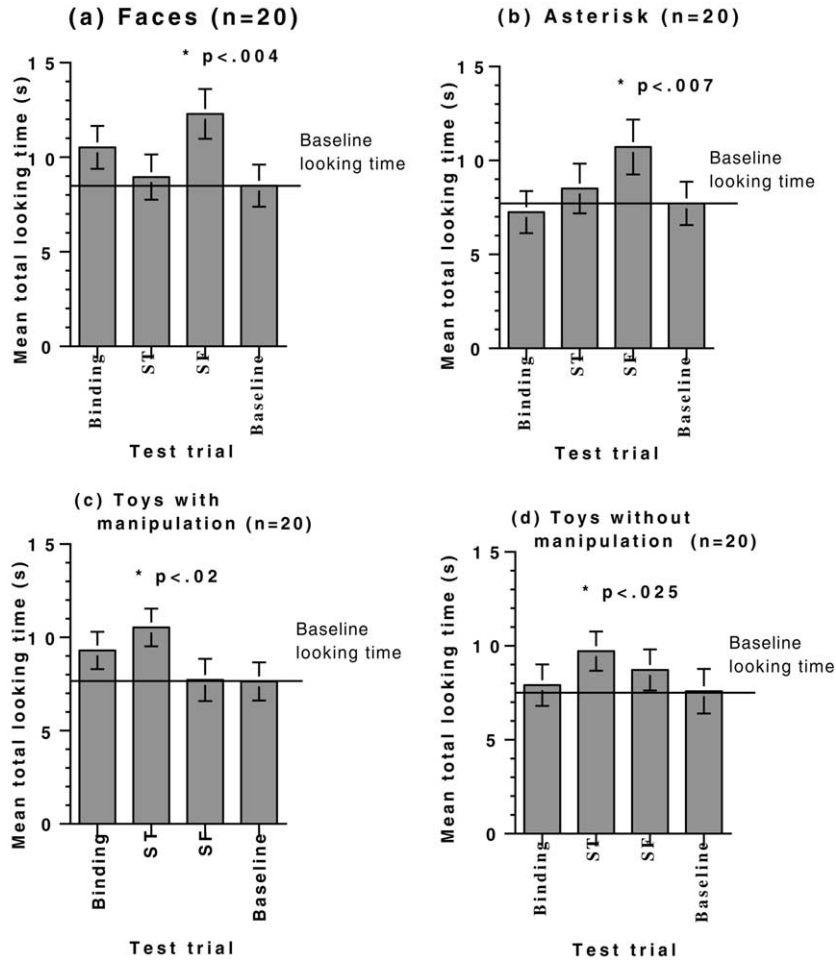


Fig. 4. Mean total looking times to test trials in each of the four experimental conditions. Standard errors are also shown.

3.2. Test trials

The mean total looking times on test trials were entered into an ANOVA with two factors: experimental Condition with four levels (Faces, Asterisks, Toys with manipulations, and Toys without manipulations) as a between-subject factor and test Trial with four levels (Binding, SF, ST, and Baseline) as a repeated measures factor. The ANOVA revealed a highly significant effect of Test trial ($F(3, 28) = 5.02$, $P < 0.003$), no significant main effect of Condition ($F(3, 76) = 0.62$, $P > 0.60$), and a highly significant Trial \times Condition interaction ($F(9, 228) = 2.79$, $P < 0.005$). These are discussed in turn below.

Fig. 3 shows the mean total looking times in the different test trials collapsed across experimental conditions (main effect of experimental Condition). The principal hypoth-

eses explored are that (1) 4-month-olds process object feature information, (2) 4-month-olds process object location information, and (3) 4-month-olds do not process the conjunction of feature and location information. These hypotheses can be assessed by comparing the looking times in the SF, ST, and Binding conditions to the looking times in the Baseline condition, respectively. Planned contrast comparisons revealed that whereas the looking times in the Binding trials do not differ significantly from Baseline ($F(1, 238) = 2.58$, $P > 0.11$), both the looking times in the SF trials ($F(1, 238) = 13.17$, $P < 0.0005$) and the ST trials ($F(1, 238) = 8.10$, $P < 0.005$) differed significantly from the looking times in the Baseline trials. In addition, a Wilcoxon ranked sign test for dependent measures revealed that while significantly more babies looked longer in the SF trials ($n = 80$, $z = 3.61$, $P < 0.001$) and the ST trials ($n = 80$, $z = 3.17$, $P < 0.002$) than they did in the Baseline trials, there were no significant differences in the number of babies who looked longer at the Binding trials as compared to the Baseline trials ($n = 80$, $z = 1.62$, $P > 0.10$). Thus, across experimental conditions, it appears that infants are responding to changes in *either* surface feature *or* location information independently, *but not to changes in the conjunction* of surface feature and location information.

However, this pattern of looking times is qualified by the interaction between experimental Condition and test Trial. To explore this interaction, we examined whether the looking times in the Binding, SF and ST test trials differed significantly from Baseline in each of the four experimental conditions separately (Fig. 4).

3.3. Performance in the Faces condition

Planned comparisons revealed that only in the SF trial did the infants look significantly longer ($F(1, 57) = 9.408$, $P < 0.003$) than in the Baseline trial. Looking times in the Binding ($F(1, 57) = 2.68$, $P > 0.11$) and ST ($F(1, 57) = 0.135$, $P > 0.68$) trials did not differ significantly from Baseline. In addition, a Wilcoxon ranked sign test revealed that more infants looked longer during the SF trials than they did in the Baseline trials ($n = 20$, $z = 2.73$, $P < 0.007$) but that there were no significant differences in the number of infants who looked longer in the ST ($n = 20$, $z = 0.30$, $P > 0.76$) or Binding trials ($n = 20$, $z = 1.53$, $P > 0.12$) than in the Baseline trials. Thus, infants in this condition appear to have encoded and remembered the feature information in the faces, but not the location information or the binding information in these test events.

This result is surprising given that previous studies have suggested that infants attended preferentially to spatial and temporal information in such occlusion events (e.g. Leslie et al., 1998; Simon, Hespos, & Rochat, 1995; Wilcox, 1999; Xu & Carey, 1996).

3.4. Performance in the Asterisks condition

Planned comparisons revealed that only in the SF trial did the infants look significantly longer ($F(1, 57) = 8.059$, $P < 0.007$) than in the Baseline trial. Looking times in the Binding ($F(1, 57) = 0.193$, $P > 0.66$) and ST ($F(1, 57) = 0.561$, $P > 0.45$) trials did not differ significantly from Baseline. In addition, a Wilcoxon ranked sign test revealed that more infants looked longer during the SF trials than they did in the Baseline trials ($n = 20$, $z = 2.95$, $P < 0.004$) but that there were no significant differences in the number

Table 2

Mean difference in looking times between test trials and the baseline test trial in seconds as a function of experimental conditions (standard errors in parentheses)

	Experimental condition			
	Ventral processing		Dorsal processing	
	Faces	Asterisks	Toys with manipulation	Toys without manipulation
Binding trial	2.03 (1.30)	– 0.47 (1.09)	1.66 (1.22)	0.32 (0.83)
SF trial	3.79 (1.23)	3.00 (0.91)	0.08 (0.96)	1.13 (1.02)
St trial	0.45 (0.94)	0.79 (0.75)	2.90 (0.93)	2.13 (0.68)

of infants who looked longer in the ST ($n = 20$, $z = 0.60$, $P > 0.54$) or Binding trials ($n = 20$, $z = 0.30$, $P > 0.76$) than in the Baseline trials. Thus, as with faces, infants in this condition appear to have encoded and remembered the color information in the asterisks, but not the location information or the binding information.

3.5. Performance in the Toys with manipulation condition

Planned comparisons revealed that only in the ST trial ($F(1, 57) = 6.004$, $P < 0.02$) did infants look significantly longer than in the Baseline trial. Looking times in the Binding ($F(1, 57) = 1.971$, $P > 0.17$) and SF ($F(1, 57) = 0.005$, $P > 0.95$) trials did not differ significantly from Baseline. In addition, a Wilcoxon ranked sign test revealed that more infants looked longer during the ST trials than they did in the Baseline trials ($n = 20$, $z = 2.60$, $P < 0.01$) but that there were no significant differences in the number of infants who looked longer in the SF ($n = 20$, $z = 0.40$, $P > 0.68$) or Binding trials ($n = 20$, $z = 1.18$, $P > 0.23$) than in the Baseline trials. Thus, infants in this condition appear to have encoded and remembered the location in the hidden toys, but not the surface feature information or the binding information.

3.6. Performance in the Toys without manipulation condition

Planned comparisons revealed that only in the ST trial did infants look significantly longer ($F(1, 57) = 5.547$, $P < 0.025$) than in the Baseline trial. Infants did not look longer than Baseline in the Binding ($F(1, 57) = 0.13$, $P > 0.62$) or SF ($F(1, 57) = 1.56$, $P > 0.21$) trials. In addition, a Wilcoxon ranked sign test revealed that more infants looked longer during the ST trial than they did in the Baseline trials ($n = 20$, $z = 2.78$, $P < 0.01$) but that there were no significant differences in the number of infants who looked longer in the SF ($n = 20$, $z = 1.20$, $P > 0.23$) or Binding trials ($n = 20$, $z = 0.11$, $P > 0.91$) than in the Baseline trials. Thus, infants in this condition appear to have encoded, and remembered the location information in the hidden toys, but not the surface feature information or the binding information.

We also asked whether the pattern of looking more than baseline varied significantly across experimental conditions. Indeed, a strong interpretation of our hypothesis suggests that, not only should infants look more than baseline to a feature violation in the first two conditions (Faces and Asterisks), but that this looking time should be greater than in the

last two conditions (Toys with and without manipulations). Conversely, not only should infants look longer than baseline toward a location violation trial in the last two experimental conditions (Toys with and without manipulation), their looking time should be greater than in the first two experimental conditions (Faces and colored Asterisks). Moreover, what predicts the cross-over in looking times is whether the target stimulus is processed by the dorsal or ventral pathways.

Table 2 shows the difference between the looking times on the different test trials and the associated baseline test trial. That is, it represents the degree to which there was an increase or decrease in looking in response to a test violation as compared to the baseline amount of looking to a test event with no change in location or feature information. A separate mixed-design ANOVA with Trial (Binding, SF, and ST) as a within-subject factor and Pathway (ventral processing, dorsal processing) as a between-subject factor was carried out on these data. The ANOVA revealed a highly significant Trial \times Pathway interaction ($F(2, 156) = 7.84, P < 0.001$). Planned contrasts revealed that there were no significant differences between the mean looking times, relative to the baseline trials, in the ventral and dorsal conditions in the binding trials ($F(1, 76) = 0.04, P > 0.85$). However, on SF trials, there was significantly more looking, relative to the baseline trials, during ventral processing conditions than during dorsal processing conditions ($F(1, 76) = 7.23, P < 0.001$), with no significant difference in looking, relative to the baseline trials, between the Faces and Asterisks conditions ($F(1, 76) = 0.29, P > 0.58$) or between the two Toy conditions ($F(1, 76) = 0.52, P > 0.47$). Similarly, in the ST trials, there was significantly more looking, relative to the baseline trials, during dorsal processing conditions than during ventral processing conditions ($F(1, 76) = 5.17, P < 0.03$), with no significant difference in looking, relative to the baseline trials, between the Face and Asterisks conditions ($F(1, 76) = 0.08, P > 0.78$) or between the two Toy conditions ($F(1, 76) = 0.42, P > 0.52$).

Finally, the lack of response to the appearance of a novel toy in the last two experimental conditions could potentially be explained by a lack of discrimination or sensitivity to the features of the different toys. However, a discrimination study revealed that this was not the case. Nine 4-month-olds (four boys and five girls with a median age of 3 months and 28 days) were habituated (using an infant-controlled habituation procedure) to a randomly selected toy image used in experimental conditions 3 and 4. When looking time had dropped to 50% of that in the previous three trials, infants were presented with a second randomly selected toy image as a test item. The average looking time duration in the test trial ($M = 5.43$ s, $SD = 0.96$) was found to be significantly greater ($t(8) = 3.04, P < 0.05$) than in the last habituation trial ($M = 2.49, SD = 1.0$), thus demonstrating the ability to discriminate between the different toy images.

To assess whether infants could discriminate between the toy images presented in the context of a complex event involving motion, we examined whether the amount of decrease in looking time that occurred during familiarization depended on the items used during familiarization. If the decrease in looking time during familiarization is independent of the objects presented then this is consistent with the proposal that the infants do not discriminate between the toy stimuli. Indeed, if they did not discriminate between items, then all items would appear identical and the presence of any individual item would not be a significant predictor of the decrease in looking time. In contrast, if the

decrease in looking time is found to depend on one, or both of the items introduced during familiarization, then this would imply that the infants can discriminate between the toy items, and that this has an impact on their behavior.

To explore these hypotheses, we carried out a series of regression analyses on the decrease in looking time between the second and final familiarization trials (Table 1). The dependent variable was the decrease in looking time, and the regressors (independent variables) were two categorical variables (object_1 and object_2) coding for the identity of which of the seven toys was the first object presented and which was the second object presented. We then rank ordered the means from smallest to largest difference and computed the regressions on the rank ordered means.⁷

Simple independent regressions showed that the identity of object_2 was a significant predictor of the decrease in looking time ($F(1, 38) = 6.03$, $P < 0.02$) and accounted for 17% of the variance. Object_1 tended towards but did not reach significance as a predictor of the decrease in looking ($F(1, 38) = 3.00$, $P < 0.10$) and accounted for 7% of the variance. A linear model with both object_1 and object_2 as independent variables was also as significant predictor of the decrease in looking time during familiarizations ($F(2, 37) = 4.19$, $P < 0.03$), accounting for 19% of the variance. As expected from the simple regression analyses, only the regression coefficient for object_2 differed significantly from 0 ($t(39) = 2.25$, $P = 0.03$), whereas the regression coefficient for object_1 did not differ significantly from 0 ($t(39) = 1.47$, $P < 0.15$).

In short, specifying the identity of the two objects presented during familiarization significantly predicts the decrease in looking time across the familiarization trials. This means that the infants' looking time behaviors depend on the particular toy images encountered, and hence, that they can and do discriminate between the toy image stimuli used in final two experimental conditions.

4. Discussion

This study examined 4-month-olds' ability to encode and maintain object information

⁷ Because the category labels are arbitrary there is no natural order among them – and certainly no guarantee of monotonicity across means as one moves through the category labels. Rank ordering the means ensures a monotonic increase in means – a requirement of linear models. This simplifies the analyses and allows us to apply the well-known linear regression methods. Our argument would still be valid without rank ordering of means, but we would need to examine the fit of non-linear functions (e.g. polynomials) rather than linear functions. Since we are only interested in showing that there is a significant relationship between what object was present and the decrease in looking time – and not what the shape of that relationship is – showing that there is a significant linear relationship among the ranked means is equivalent to showing that there is a significant non-linear relationship among the unranked means. However, to ensure that there was a true significant difference in looking time drop-off as a function of which object was used in familiarization, *prior to the rank ordering of means*, we compared the mean drop-off for every object level of both object_1 and object_2 on the unranked data. This revealed a marginally significant difference between the lowest and highest means (the end points of the rank ordering) obtained for the different levels of object_1 ($t(39) = -1.98$, $P < 0.06$, two-way, Dunnett multiple comparison procedure), and a significant difference for the different levels of object_2 ($t(39) = 2.18$, $P < 0.04$, two-way, Dunnett multiple comparison procedure). This very conservative analysis echoes the regression results on the ranked means in which object_1 was only a marginally significant predictor of looking times whereas object_2 was a significant predictor of looking times.

following brief occlusions. Colors, faces, location, and “potentially manipulable” were selected as target traits because they are assumed to be diagnostic of ventral or dorsal route processing (e.g. Milner & Goodale, 1995). When the occluded objects were female faces, the 4-month-olds responded only to violations of surface feature information and did not respond to novel spatial information or the conjunction of identity and location information following occlusion. This finding stands in contrast to previous work suggesting that young infants encode and maintain spatial and temporal information preferentially in similar occlusion tasks (Kaldy & Leslie, *in press*; Leslie et al., 1998; Simon et al., 1995; Wilcox, 1999; Xu, 1999; Xu & Carey, 1996). The pattern of looking times obtained with monochromatic asterisks as target objects was the same as that obtained with faces. The infants responded only to novel surface feature information (color) but not to novel location information or the novel conjunction of feature and location information, thereby ruling out the possibility that the preferential encoding and maintenance of surface feature information was unique to face stimuli. In contrast, when presented with images of toys as target objects, the infants responded only to the novel location information but not the novel surface feature or the novel conjunction of location and surface feature information. Moreover, this was true whether or not the infants had seen the toys being manipulated immediately prior to testing.

Because the events are all computer controlled, the duration, timing and motion of all events are identical in all experimental conditions. Thus, differences in looking times cannot be attributed to differences to any of these variables. In addition, there is an equivalent decrease in looking times during familiarization in all conditions, suggesting that there is an equivalent level of encoding occurring in all conditions so the different response patterns in the different experimental conditions can not be attributed to differing degrees of familiarity with the test stimuli.

Perhaps the infants have not even encoded the neglected information in the different experimental conditions? However, because in the present study the kind information that is maintained (location or identity) is target dependent, it follows that even in the case of toys where only location information was maintained, feature information had to be processed in order to establish what information to discard and what information to maintain.

Tremoulet et al. (2000) found that somewhat older infants do show increased looking to the appearance of a novel location and shape conjunction following brief occlusions. This suggests that our results cannot simply be attributed to an inability of infants to remember more than one dimension of an occluded stimulus. There must be some additional factor that limits the ability of 4-month-olds to maintain the conjunction of color or identity with location, but that does not apply to the conjunction of shape and location information.

So why did the infants in this study respond so differently from infants in other, similar studies? Both color and faces are diagnostic markers of ventral stream processing. Milner and Goodale (1995) suggested that the ventral stream develops representations used for recognition of objects whereas the dorsal stream develops representations used for acting on objects. Studies that find an early reliance on spatial and temporal information have tested infants with real three-dimensional objects that can potentially be acted upon, and are often acted on by the experimenter. Unlike real objects, the target stimuli in the first two experimental conditions (two-dimensional disembodied faces and monochromatic

asterisks) do not naturally afford action. We believe that this is the basis for the selective maintenance of object information during occlusion.

We are not claiming that information is maintained by the dorsal or ventral stream, only that the information processed by the dorsal and ventral streams is maintained (somewhere/somehow). This presumably involves the frontal lobes (Rao et al., 1997), which are not fully mature at this age, hence the inability to maintain both sets of information during the occlusion.

These results are consistent with the proposal made by Mareschal et al. (1999) and Johnson et al. (2001) that young infants have difficulty integrating information processed independently in the dorsal and ventral streams. In the present study, we have suggested that the affordance of the object was the key feature determining which of the two representations is most likely to be maintained during an occlusion. However, although the model predictions are consistent with the overall Condition effect depicted in Fig. 3 (i.e. the independent processing of dorsal and ventral information), it does not predict the selective maintenance of one kind of information over the other. In addition, the model is not sensitive to object affordances.

These results have general implications for the design of studies aiming to probe infant knowledge of objects. They suggest that because there might be differential processing of the stimulus features depending on the kind of stimulus used, great care needs to be taken both in selecting the stimuli used, and also in interpreting a failure of infants to use a particular type of stimulus information. Failures to demonstrate the use of information in a subsequent test phase may not reflect a lack of processing, but rather, reflect the outcome of competition in which information judged less vital is not retained. For example, while Xu and Carey (1996) suggested that 10-month-olds did not use surface feature information to individuate objects when tested with real three-dimensional toys, Bonatti, Frot, Zangl, and Mehler (2002) have recently found that 10-month-olds will process feature information if tested with face stimuli similar to those used in the paper.

The fact that 4-month-olds were sensitive to the affordance of the targets raises an interesting question about the ontogeny of visual affordances. Tucker and Ellis (2001) argue that these visual-motor associations reflect a participant's previous experience and interactions with objects. However, this cannot be the case with our infants because reaching towards an object begins at about 4.5 months (Gordan, 2001). So, how could 4-month-old infants have acquired a sensitivity to action affordances in objects?

One possibility is through simple observation. Although 4-month-olds do not reliably reach yet, they are on the cusp of reaching. The toys photographed for use in this experiment were common, commercially available infant toys. They may have recognized the objects as members of the category of objects that they routinely encounter and observe being used in particular ways. Certainly, adults can categorize visually presented objects on the basis of the actions that they afford (Klatzky, Pellegrino, McCloskey, & Lederman, 1993; Tucker & Ellis, 2001). This suggests that infants may initially acquire crude object–action associations through observation alone. These initial associations may underpin early efforts to engage in purposeful action with an object. Clearly, further work is required to tease these possibilities apart.

In summary, 4-month-olds have a limited ability to maintain at the same time object information processed separately in the dorsal and ventral visual streams. This results in

the selective loss of information processed in one or the other stream following occlusion events. Furthermore, our findings suggest that it is the target's affordance for action that determines whether the dorsal or ventral information is selectively maintained.

Acknowledgements

We are very grateful to Eileen Mansfield and Seok Hui Tan for helping to collect the data and to Daisy Powell for helping to score the data. The image bank from which the female faces were drawn was put together by Daphne Maurer and Sybil Geldart from McMaster University, Canada. We acknowledge financial support from the Royal Society, the Sackler Institute, Economic and Social Research Council grant R000239112, European Commission RTN grant HPRN-CT-2000-00065, and Medical Research Council Programme Grant (G97 15587).

References

- Baillargeon, R. (1993). The object concept revisited: new directions in the investigation of infants' physical knowledge. In C. E. Granrud (Ed.), *Visual perception and cognition in infancy* (pp. 265–315). London: Lawrence Erlbaum Associates.
- Bonatti, L., Frot, E., Zangl, R., & Mehler, J. (2002). The human first hypothesis: identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology*, 44, 388–426.
- Craighero, L., Fadiga, L., Rizzolatti, G., & Umiltà, C. (1999). Action for perception: a motor-visual attentional effect. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1673–1692.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature*, 349, 154–156.
- Gordan, A. M. (2001). Development of hand motor control. In A. F. Kalverboer & A. Gramsbergen (Eds.), *Handbook of brain and behaviour in human development* (pp. 513–538). Dordrecht: Kluwer Academic Press.
- Grèze, J., & Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia*, 40, 212–222.
- Gross, C. G. (1992). Representation of visual stimuli in inferior temporal cortex. *Philosophical Transactions of the Royal Society of London*, B335, 3–110.
- Haaf, R. A., Brewster, M., de Saint-Victor, C. M., & Smith, P. H. (1989). Observer accuracy and observer agreement in measurement of visual fixation with fixed trial procedures. *Infant Behavior and Development*, 12, 211–230.
- Jeannerod, M. (1988). *The neural and behavioural organization of goal-directed movements*, Oxford: Oxford University Press.
- Johnson, M. H., Mareschal, D., & Csibra, G. (2001). The functional development and integration of the dorsal and ventral visual pathways: a neurocomputational approach. In C. A. Nelson & M. Luciana (Eds.), *The handbook of developmental cognitive neuroscience* (pp. 339–351). Cambridge, MA: MIT Press.
- Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development*. Oxford: Blackwell.
- Kaldy, Z., & Leslie, A. M. (in press). Identification of objects in 9 month old infants: integrating what and where information. *Developmental Science*.
- Klatzky, R. L., Pellegrino, J. W., McCloskey, B. P., & Lederman, S. J. (1993). Cognitive representations of functional interactions with objects. *Memory and Cognition*, 21, 253–270.
- Komatsu, H., Ideura, Y., Kaji, S., & Yamane, S. (1992). Color selectivity of neurons in the inferior temporal cortex of the awake macaque monkey. *Journal of Neuroscience*, 12, 408–424.
- Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: developing 'what' and 'where' systems. *Trends in Cognitive Sciences*, 2, 10–18.

- Mareschal, D., Plunkett, K., & Harris, P. (1999). A computational and neuropsychological account of object-oriented behaviors in infancy. *Developmental Science*, 2, 306–317.
- Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.
- Newsome, W. T., Wurtz, R. H., & Komatsu, H. (1988). Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. *Journal of Neurophysiology*, 60, 604–620.
- Perret, D. I., Mistlin, A. J., & Chitty, A. J. (1987). Visual neurones responsive to faces. *Trends in Neurosciences*, 10, 358–364.
- Rao, S. C., Rainer, G., & Miller, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, 276, 821–824.
- Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. In C. Umiltà & M. Moscovitch (Eds.), *Conscious and nonconscious information processing: Attention and performance* (pp. 231–265), Vol. XV. Cambridge, MA: MIT Press.
- Simon, T. J., Hespos, S. J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn. *Cognitive Development*, 10, 253–269.
- Spelke, S. E., Katz, G., Purcell, S. E., Ehrlich, S. M., & Breilinger, K. (1994). Early knowledge of object motion: continuity and inertia. *Cognition*, 51, 131–176.
- Taira, M., Mine, S., Georgopoulos, A. P., Mutara, A., & Sakata, H. (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movements. *Experimental Brain Research*, 83, 29–36.
- Tremoulet, P. D., Leslie, A. M., & Hall, D. G. (2000). Infant individuation and identification of objects. *Cognitive Development*, 15, 499–522.
- Tucker, M., & Ellis, R. (1998). On the relation of seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 830–846.
- Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition*, 8, 769–800.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.
- Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: an integrated systems perspective. *Science*, 255, 419–423.
- Wilcox, T. (1999). Object individuation: infants' use of shape, size, pattern, and color. *Cognition*, 72, 125–166.
- Wilcox, T., & Baillargeon, R. (1998). Object individuation in infancy: the use of featural information in reasoning about occlusion events. *Cognitive Psychology*, 17, 97–155.
- Wilcox, T., & Schweinle, A. (2002). Object individuation and event mapping: developmental changes in infants' use of featural information. *Developmental Science*, 5, 132–150.
- Wynn, K. (1992). Addition and subtraction in human infants. *Nature*, 358, 749–750.
- Xu, F. (1999). Object individuation and object identity in infancy: the role of spatiotemporal information, object property information, and language. *Acta Psychologica*, 102, 113–136.
- Xu, F., & Carey, S. (1996). Infants' metaphysics: the case of numerical identity. *Cognitive Psychology*, 30, 111–153.
- Xu, F., & Carey, S. (2000). The emergence of kind concepts: a rejoinder to Needham and Baillargeon (2000). *Cognition*, 74, 285–301.
- Xu, F., Carey, S., & Welch, J. (1999). Infants' ability to use object kind information for object individuation. *Cognition*, 70, 255–284.