Cognitive Development, 11, 161–180 (1996)

Perception of Object Unity in Young Infants: The Roles of Motion, Depth, and Orientation

Scott P. Johnson
Richard N. Aslin
University of Rochester

One hundred twenty-eight 4-month-old infants were habituated to one of several displays that depicted two rod pieces above and below a box. The effects of common motion, background texture, and orientation of the rod pieces on infants' perception of unity of the partially occluded rod were examined. Infants who viewed displays in which the rod pieces were aligned and presented in front of a textured background, subsequently looked longer at a broken rod (two rod pieces separated by a gap) than at a complete rod, implying that the infants experienced the rod pieces as connected behind the box in the first display. Infants who viewed displays with no background texture, or displays in which the rod pieces were nonaligned but relatable (i.e., connected if extended behind the occluder), looked equally at the two posthabituation displays. Infants who viewed displays containing nonrelatable rod pieces looked longer at the complete rod, implying that nonrelatable edges specify disjoint objects to 4-month-olds. A threshold model, stipulating that perception of object unity is supported by multiple visual cues, is proposed to account for these results. Veridical perception of motion of display elements, depth ordering, and edge orientation are necessary, but not individually sufficient, to support young infants' perception of object unity.

The environment is filled with objects. Information for the layout of objects is available to the observer by various visual cues such as color, luminance, contour, texture, depth, and orientation of surfaces and edges. The observer's task is to parse the visual array into entities that are distinct and segregated, on the basis of these visual cues. This must be accomplished in spite of potentially formidable challenges to the process. For example, ob-

---

This research was supported by NIH postdoctoral training grants EY-07125 and HD-07748 to SPJ, NSF grant SBR-9108723 to RNA, and NEI grant EY-01319 to the Center for Visual Science, University of Rochester. Portions of this research were presented at the June 1994 International Conference on Infant Studies, Paris. The authors wish to thank Elizabeth Spelke for helpful comments, Jill Gallipeau and Sara Matteson for serving as observers, and especially the infants and parents who participated in the studies.

Correspondence and requests for reprints should be sent to Scott P. Johnson, Department of Psychology, Lancaster University, Lancaster, LA1 4YF, England. E-mail: <S.P.Johnson@lancaster.ac.uk.>.

jects move into and out of sight, or may be partly or completely occluded by other objects, or two objects may be adjacent.

How do infants organize visual scenes? One way in which the question of the development of sensitivity to visual information has been addressed is by consideration of Gestalt phenomena (e.g., Spelke, 1990). Gestalt psychologists espoused the view that humans tend to organize visual scenes according to principles such as maximization of the featural goodness of a scene, and by analyzing scenes in terms of good form, common fate, similarity, and good continuation of surfaces and edges in the array (Koffka, 1935).

Previous research has tended to favor the hypothesis that young infants are not proficient at using many Gestalt cues in perception of object layout. For example, Spelke, Breinlinger, Jacobson, and Phillips (1993) found that at 9 months of age, infants were only weakly sensitive to similarity, good form, and good continuation in perceiving objects. Likewise, Kellman and Spelke (1983) reported that 4-month-olds did not respond to similarity, good form, and good continuation, but were sensitive to common fate in perception of partly occluded objects.

Kellman and Spelke (1983) employed displays that depicted a partially occluded rod undergoing lateral motion behind an occluding box (see Figure 1). To most adults, the display in Figure 1 gives rise to an unambiguous percept of a single, partially hidden object behind a surface. This percept in adults generalizes to cases in which display elements are stationary (Kellman & Spelke, 1983), or two-dimensional (Johnson & Náñez, 1995). A variety of studies have demonstrated that young infants also perceive the partially occluded rod as a single object extending behind the box (Johnson & Aslin, 1995; Johnson & Náñez, 1995; Kellman & Spelke, 1983; Kellman, Spelke, & Short, 1986; Slater et al., 1990).

Perception of object unity in these studies was assessed with a habituation paradigm. Infants were first shown two rod pieces, above and below an occluding box, until habituation of looking occurred. The infants then viewed test displays consisting of a complete rod, or two rod pieces with a

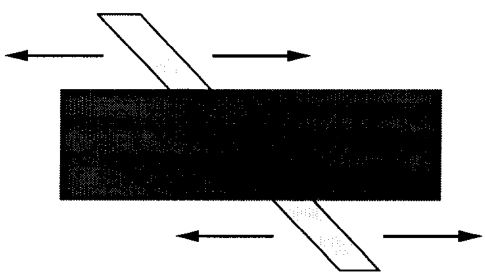


Figure 1. Schematic depiction of an occlusion display typically used in object unity studies.

visible gap corresponding to the location of the box in the habituation displays (broken rod). The test display that elicits the most interest is inferred to be more novel relative to the habituation display. In general, preference for the broken rod test display indicates perception of unity of the rod pieces in the rod-and-box habituation display. A lack of a clear preference may indicate an ambiguous percept of the unity of the rod pieces, and a preference for the complete rod may indicate that the rod pieces were perceived as disjoint objects.

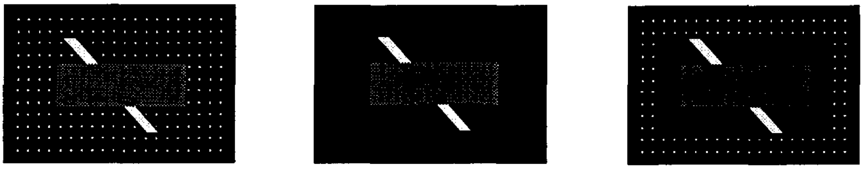
Young infants do not seem to perceive the unity of the rod pieces in such displays under some conditions. For example, perception of object unity in motion displays has not been demonstrated before 2 months of age (Johnson & Aslin, 1995; Slater, Johnson, Brown, & Badenoch, 1996; Slater, Johnson, Kellman, & Spelke, 1994). Perception of object unity in stationary rod-and-box displays has not been observed earlier than 5 months of age (Craton, 1994). The studies presented here were intended to extend our knowledge of the visual information used by young infants in the perception of object unity.

Kellman (1993; Kellman & Shipley, 1991) proposed a two-process theory of perceptual unit formation (by unit formation we mean assigning two disparate surfaces to a single, partly occluded object). The first process, the *primitive* process, takes into account common motion of visible surfaces in determining their unity. The primitive process is insensitive to the orientations of edges that lead behind an occluder. The second process, the *rich* process, takes into account edge orientations as well as their motions. Edges on either side of the occluder will be perceived as belonging to a single edge if they are *relatable* when interpolated (perceptually extended) behind the occluder. Relatability is defined mathematically, but is based on the requirement that disparate edges will be perceived as unified if they can be connected with a smooth, monotonic curve behind the occluder (see Kellman & Shipley, 1991, for details).

According to Kellman (1993), only the primitive process is operational in infants younger than 6 months of age. Thus the two-process theory predicts that two disparate surfaces that extend from behind an occluder will be perceived as unified by 4-month-olds if the surfaces undergo common motion, regardless of their orientations or depth relations between the surfaces. The following experiments tested this prediction. All participants were 4 months of age, a group in which perception of object unity has been demonstrated to be robust (Johnson & Náñez, 1995; Kellman & Spelke, 1983; Slater et al., 1990).

EXPERIMENT 1

Experiment 1 employed rod-and-box displays in which the two visible parts of the rod were aligned (and thus relatable), and underwent common lateral



(A) (B) (C)

Figure 2. The displays used in Experiment 1. (A) Texture, aligned, relatable (T/A/R) habituation display, (B) No texture, aligned, relatable (NT/A/R) habituation display, and (C) Surrounding texture, aligned, relatable (ST/A/R) habituation display.

translation. As seen in Figure 2a, the T/A/R stimulus consisted of two rod pieces above and below a box, against a textured background (consisting of a grid of white dots). The label T/A/R denotes the presence of background texture, as well as alignment and relatability of the edges of the rod pieces. As seen in Figure 2b, the NT/A/R stimulus was identical to the T/A/R stimulus, except there was no texture (the background was black). We predicted that if common motion is sufficient to specify the unity of the rod pieces to 4-month-olds, then infants who were habituated to either the T/A/R or the NT/A/R display should look longer subsequently at a broken rod than at a complete rod. On the other hand, if accretion and deletion of background texture serves to aid in the process of young infants' unit formation (perhaps as an additional depth cue), then only infants in the T/A/R group should look longer at a broken rod test display. (The T/A/R display was similar to that employed by Johnson & Náñez, 1995, who found evidence of perception of object unity by 4-month-olds in such a display.)

Method

Participants. Sixty-four full-term infants (35 girls) comprised the final sample (mean age = 126 days, range = 106–138 days). An additional 9 infants were observed but not included in the sample due to excessive fussiness (8 infants) or sleepiness (1 infant). The infants were recruited by letter and telephone from hospital records and birth announcements in the local newspaper. The majority of the participants were from white, middle-class families. Parents were paid $5.00 for their participation.

Apparatus and Stimuli. An Amiga 3000 computer and an 84 cm Sony color monitor were used to generate the displays. Two observers viewed the infant through small peepholes cut into either side of a black panel that extended 47 cm from the sides of the monitor (see Figure 3).

The computer presented the stimulus displays, stored each observer's data, calculated the habituation criterion for each infant, and changed dis-

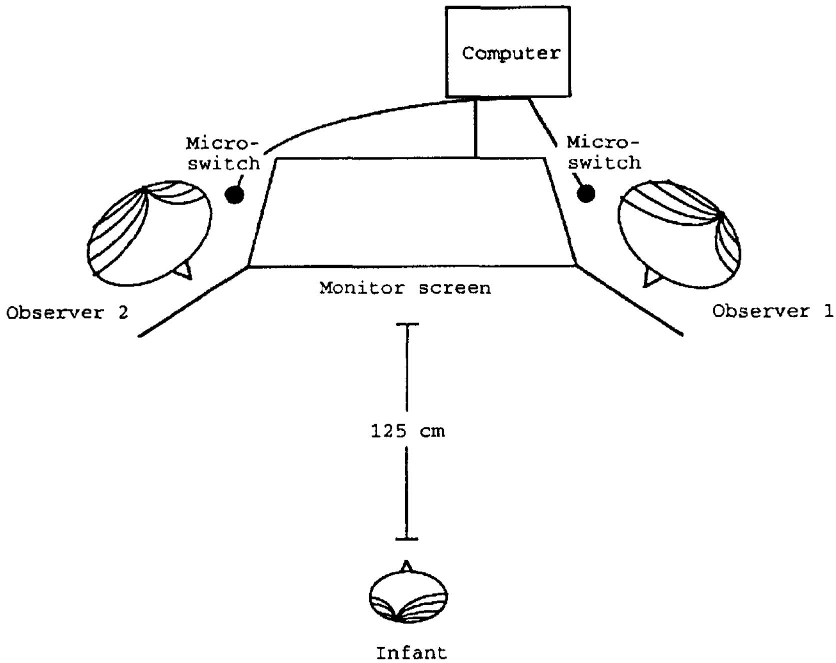


Figure 3. Schematic overhead view of the experimental setup.

plays after the criterion was met. The computer also recorded how long the infant looked at each display, according to the observers' judgments. These judgments were entered via two hand-held microswitches, connected to the computer's mouse port. Both observers were blind to the stimulus on the screen at any given time. The second observer had never been allowed to view the displays, and was naive to the hypotheses under investigation.

The T/A/R display consisted of a computer-generated 33 × 12.7 cm blue box, subtending 15 × 5.8° visual angle (at the infants' 125 cm viewing distance). The box was oriented with its long axis horizontal. A yellow rod, 33 cm in length (15°) and oriented 42° counterclockwise from the vertical, underwent lateral translation at a rate of 10.5 cm/s (4.8°/s) behind the box. The center portion of the rod appeared to be occluded by the box. The rod and box were presented against a textured background, consisting of a regular 12 × 20 grid of white dots (see Figure 2a).

The two test stimuli were similar to the rod portion of the habituation stimulus, but without the box. The broken rod contained a 12.7 cm (5.8°) gap in its center. The broken rod and complete rod test displays moved in the same translatory motion as the two rod pieces in the T/A/R display, against a textured background (the background texture was visible in the gap of the broken rod). The test displays were presented singly, in alternation.

The NT/A/R display was identical to the T/A/R display, except there were no dots as background texture. The two test stimuli that followed habituation to the NT/A/R display likewise were identical to those that followed the T/A/R display, but without background texture.

A control display was employed to examine a possible inherent preference for either of the two test displays. The NT/A/R control display was identical to the NT/A/R display, except the bottom rod piece remained stationary as the top piece moved laterally.[1] The two test displays that followed were identical to those presented after the NT/A/R display. (Johnson & Náñez, 1995, found no preference for either test display after habituation to a T/A/R control display.)

Sixteen infants were habituated to each display, either the T/A/R, NT/A/R, or NT/A/R control display. (An additional 16 infants were presented with a fourth display, labeled ST/A/R, to address a question that arose from the results of the first three displays. See Results and Discussion section for details.) In each group of 16 infants, 8 viewed the broken rod first after habituation, and 8 viewed the complete rod first (order was determined randomly by the computer).

Procedure. The infants were placed in an infant seat approximately 125 cm from the display monitor. The display was presented until each infant met the habituation criterion. This criterion was defined according to a common infant-control procedure (Horowitz, Paden, Bhana, & Self, 1972) as a decline in looking time during three consecutive trials, adding up to less than half the total looking time during the first three trials. If the total looking time during the first three trials was less than 12, the criterion was based on the first three subsequent trials for which looking time totaled 12 or more. The habituation period was terminated if an infant had not met the criterion after 15 trials ($N = 4$ for both Experiments 1 and 2). The test period followed as with the other infants.

Timing of each trial began when the infant fixated the screen after display onset. Each observer independently indicated how long the infant looked at the display by pressing a separate microswitch as long as the infant fixated the screen, and releasing it when the infant looked away. An individual trial was terminated when both observers released their microswitches

[1]This kind of control display has been criticized because it is assumed to support a percept of a disjoint rod; thus infants habituated to this type of display should reliably dishabituate to a complete rod. However, infants habituated to such a display have been found to show no posthabituation preference (Johnson & Aslin, 1995; Johnson & Náñez, 1995; Kellman & Spelke, 1983). Moreover, infants habituated to rod-and-box displays in which the top and bottom rod pieces undergo out-of-phase motion also show no posthabituation preference (Johnson, 1995). Therefore, this type of display has strong empirical support for efficacious control of posthabituation preferences.

for two overlapping seconds. At this point, the screen was turned off by the computer, and the next display appeared two seconds later.

When looking times to the habituation display declined to criterion, the computer changed from habituation to test displays. The two test displays were seen three times each, for a total of six posthabituation trials.

Results and Discussion

Each infant contributed six posthabituation looking times to the analyses, three for the broken rod and three for the complete rod. (One infant in the NT/A/R control condition viewed only one pair of test displays after habituation, and one infant in the ST/A/R condition viewed only two pairs, due to fussiness.) Looking times were calculated by averaging the two observers' judgments for each test trial. Interobserver agreement was high for the infants included in the analyses (Pearson *r*'s averaged .98, range = .82–.99).

Before entering looking times into the analyses, they were examined for outliers. These were scores that seemed unusually long, perhaps due to difficulty in disengaging attention on the infant's part (so-called "obligatory attention"; see Johnson, 1990). These extreme scores may not be indicative of interest in the displays. An outlier was defined as a score that exceeded three *SD*'s from the mean for its cell, and was replaced by the cell mean. There were 2 outliers, out of a total of 378 scores (there were 5 outliers in Experiment 2).

Figures 4a, 4b, and 4c show the average looking times for infants who viewed the T/A/R, NT/A/R, and NT/A/R control displays, respectively. Only those infants habituated to the T/A/R display consistently looked longer at the broken rod than at the complete rod.

This observation was confirmed by analyses of looking times and patterns of dishabituation. A comparison of looking times to the broken and complete rod test displays was conducted with a 2 (group: T/A/R vs. NT/A/R) × 2 (display: broken rod vs. complete rod) × 2 (order: broken rod vs. complete rod first after habituation) × 3 (trial: first, second, or third block of test trials) repeated-measures MANOVA.

There was a significant effect of group, $F(1, 28) = 7.05, p < .05$, due to greater looking overall by infants in the T/A/R group. There was a significant effect of display, $F(1, 28) = 9.81, p < .01$, resulting from greater looking overall at the broken rod than at the complete rod. There was a significant effect of trial, $F(2, 56) = 5.09, p < .05$, the result of a decline in looking across the three blocks of test trials. There was also a significant interaction between group and display, $F(1, 42) = 8.54, p < .01$.

The interaction between group and display was due to significantly longer looking at the broken rod ($M = 14.53$ s, $SD = 16.13$) than at the complete rod ($M = 7.13$ s, $SD = 6.10$) by infants habituated to the T/A/R display, $F(1, 14) = 11.35, p < .01$. In contrast, infants habituated to the

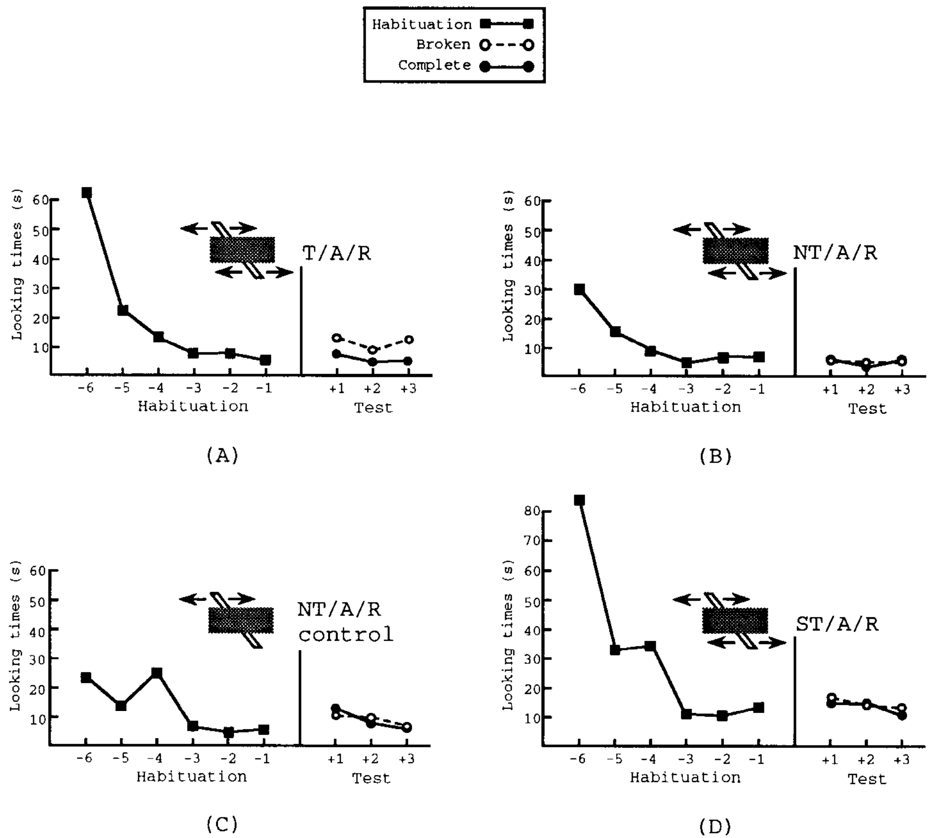


Figure 4. Mean looking times by infants in (A) the T/A/R condition, (B) the NT/A/R condition, (C) the NT/A/R control condition, and (D) the ST/A/R condition. Only the infants habituated to the T/A/R display consistently dishabituated to the broken rod, while remaining habituated to the complete rod.

NT/A/R display looked about equally at the broken rod (M = 4.56 s, SD = 4.61) and at the complete rod (M = 4.93 s, SD = 4.38), $F(1, 14) = .10$.

Tests of dishabituation to the broken and complete rod displays were conducted by comparing looking times during the last habituation trial to the average of the three looking times to each of the test displays. Infants habituated to the T/A/R display showed significant dishabituation to the broken rod, $t(15) = 3.32, p < .01$, but not to the complete rod, $t(15) = 1.90$, *ns*. In contrast, infants habituated to the NT/A/R display did not dishabituate to either the broken rod, $t(15) = .61$, or to the complete rod, $t(15) = .50$.

There were no significant main effects or interactions resulting from an analysis of looking times in the NT/A/R control group, nor did these infants dishabituate either to the broken rod, $t(15) = 1.19$, *ns*, or the complete rod, $t(15) = 1.56$, *ns*.

These results suggest that 4-month-olds seem to perceive object unity in two-dimensional displays containing accretion and deletion of texture, but not in displays without background texture. Apparently the visual information in support of unit formation remaining in the NT/A/R display (interposition, alignment, relatability, and common motion of rod pieces) is insufficient to specify object unity to 4-month-old infants. However, it may be that the infants who were habituated to the NT/A/R displays did not appear to perceive object unity not because of insufficient visual information, but because of a more general lack of interest in the displays. Note that looking times overall to the test displays were significantly lower for the NT/A/R group than the T/A/R group. If the infants in the NT/A/R group did not attend to the displays for a sufficient length of time, they may not have encoded the information for unit formation in the same way as did the infants in the T/A/R group.

To address this possibility, an additional 16 infants were habituated to the ST/A/R display, for surrounding texture/aligned/relatable (see Figure 2c). In this display, a double row of dots surrounded the rod and box, but there were no dots behind the rod. The broken rod and complete rod test displays also contained surrounding texture. It was hypothesized that the extra dots would serve to increase general interest in the display, without providing additional depth information.

As can be seen in Figure 4d, looking times overall to the test displays were higher for the ST/A/R group than for the NT/A/R group ($F(1, 27) = 29.94, p < .001$), but there was no preference for the broken rod over the complete rod after habituation to the ST/A/R display, $F(1, 13) = .47$, *ns*. These infants did not dishabituate to either the broken rod, $t(15) = .43$, or the complete rod, $t(15) = .62$. Thus the addition of surrounding texture had the effect of increasing overall attention to the test displays, but did not provide additional depth information that could contribute to the perception of the rod pieces as part of a single object, distinct from the box.

EXPERIMENT 2

Experiment 1 demonstrated that the unit formation process in 4-month-olds is not necessarily driven by common motion of partly occluded surfaces. Experiment 2 provided a more direct test of Kellman's (1993) theory, by presenting 4-month-olds with rod-and-box displays in which the edges of the rod were nonaligned and relatable, or nonaligned and nonrelatable (see Figure 5).

In Figure 5a, the edges of the rod pieces, if interpolated, meet behind the box at an angle of approximately 138°. That is, the edges are not aligned, but they are relatable according to Kellman and Shipley's (1991) criteria. This display is termed T/NA/R (texture/nonaligned/relatable). In Figure 5d, the edges of the rod pieces would not meet behind the box if interpolated. This display is termed T/NA/NR (texture/nonaligned/nonrelatable). In both the T/NA/R and T/NA/NR displays, the top and bottom rod pieces underwent common lateral translation.

It was hypothesized that if common motion of the rod parts, along with interposition and texture, are sufficient to specify object unity to 4-month-olds, then infants habituated to the T/NA/R or T/NA/NR displays should look longer at the broken rod test displays, Figures 5c and 5f, than at the complete rod displays, Figures 5b and 5e, respectively. On the other hand, if the unit formation process depends on alignment and/or relatability of edges for young infants, then perception of object unity in T/NA/R and

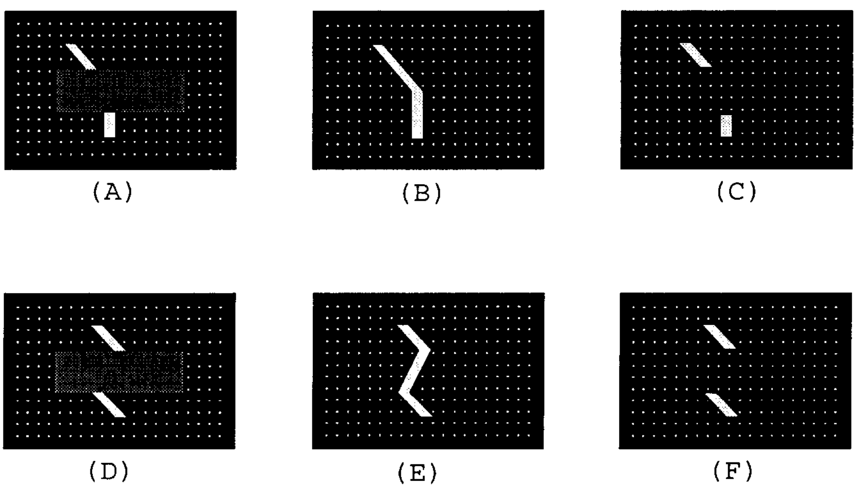


Figure 5. The displays used in the Experiment 2. (A) Texture, nonaligned, relatable (T/NA/R) habituation display, (B) T/NA/R complete rod display, (C) T/NA/R broken rod display, (D) Texture, nonaligned, nonrelatable (T/NA/NR) habituation display, (E) T/NA/NR complete rod display, and (F) T/NA/NR broken rod display.

T/NA/NR displays might be disrupted. In this case, the infants in Experiment 2 might demonstrate no posthabituation preference, or they might prefer the complete rod.

Method

Participants. Sixty-four full-term infants (26 girls) comprised the final sample (mean age = 125 days, range = 108–144 days). An additional 4 infants were observed but not included in the sample due to fussiness (2) or low interrater agreement (2) (Pearson $r < .80$). The infants were recruited from the same subject pool as those in Experiment 1.

Apparatus and Procedure. The apparatus and procedure were the same as those in Experiment 1.

Stimuli. The T/NA/R display was similar to the T/A/R rod-and-box display, except the bottom rod piece was oriented vertically. The T/NA/NR display was also similar to the T/A/R display, except for lateral displacement of the bottom rod piece, such that it was directly below the top piece. In both the T/NA/R and T/NA/NR displays, there was a 12 × 20 grid of regularly spaced white dots serving as background texture.

The test stimuli were similar to the rod portions of the T/NA/R and T/NA/NR displays, without the box. The broken rods contained a 12.7 cm (5.8°) gap in their centers. The complete and broken rods moved in the same translatory motion as the rods in the habituation displays, at 10.5 cm/s (4.8°/s). Note that the complete rod for the T/NA/R display was bent at an angle of 138°, and that the complete rod for the T/NA/NR display contained three rod segments; that is the top and bottom rod pieces were joined by a third rod piece to form a single surface.

In both the T/NA/R and T/NA/NR control displays, the bottom rod piece remained stationary whereas the top piece moved laterally. The test displays that followed habituation were identical to those presented after the T/NA/R and T/NA/NR displays, respectively.

Results and Discussion

As in Experiment 1, each infant contributed six posthabituation looking times to the analyses, three for the broken rod and three for the complete rod. (One infant in the T/NA/R condition viewed only two pairs of test displays, and one infant in the T/NA/R control condition viewed only one pair, due to fussiness.) Interobserver agreement (Pearson r) averaged .98, range = .88 to .99.

Figure 6a shows the average looking times for infants who viewed the T/NA/R display, and Figure 6b shows looking times for the T/NA/R control

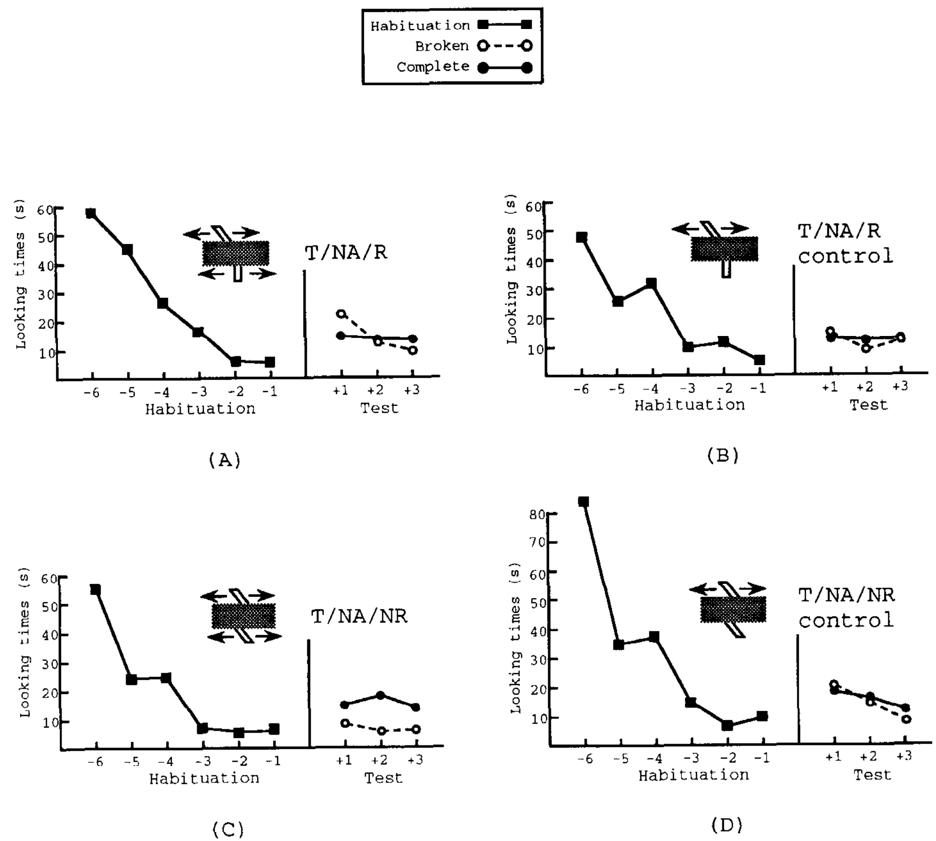


Figure 6. Mean looking times by infants in (A) the T/NA/R condition, (B) the T/NA/R control condition, (C) the T/NA/NR condition, and (D) the T/NA/NR control condition. The infants habituated to the T/NA/NR display consistently dishabituated to the complete rod, while remaining habituated to the broken rod.

display. Infants in both conditions looked about equally at the two test displays. Figure 6c shows looking times for infants who viewed the T/NA/NR display. These infants looked longer at the *complete* rod than at the broken rod. Infants who viewed the T/NA/NR control display (Figure 6d) looked about equally at the two test displays.

Comparisons of looking times to the broken and complete rod test displays were conducted separately for the T/NA/R and T/NA/NR groups with 2 (condition: experimental vs. control) × 2 (display: broken rod vs. complete rod) × 2 (order: broken rod vs. complete rod first after habituation) × 3 (trial: first, second, or third block of test trials) repeated-measures MANOVAs. Tests of dishabituation to the broken and complete rod displays were also conducted.

There were no significant main effects or interactions for the T/NA/R group, except an interaction between condition, display, and order, $F(1, 28) = 4.33, p < .05$ (this interaction was difficult to interpret, as there were no significant interactions between display and order for either the T/NA/R or the T/NA/R control conditions). There were no significant main effects or interactions resulting from display × order × trial MANOVAs run separately on the experimental and control conditions. Infants habituated to the T/NA/R display dishabituated both to the broken rod, $t(15) = 3.11, p < .01$, and to the complete rod, $t(15) = 2.42, p < .05$. Likewise, infants habituated to the T/NA/R control display dishabituated both to the broken rod, $t(15) = 3.23, p < .01$, and to the complete rod, $t(15) = 3.44, p < .01$.

There were no significant main effects or interactions for the T/NA/NR group. There was a marginally significant interaction between condition and display, $F(1, 26) = 2.92, p = .10$. Infants habituated to the T/NA/NR display looked longer at the *complete* rod, ($M = 15.60$ s, $SD = 18.74$) than at the broken rod ($M = 8.61$ s, $SD = 10.98$), $F(1, 13) = 3.50, p = .08$. Note that this result is in the direction opposite to the pattern expected if the infants perceived object unity. Infants habituated to the T/NA/NR control display did not look longer at either the broken rod ($M = 14.74$ s, $SD = 21.81$), or the complete rod, ($M = 15.73$ s, $SD = 16.51$), $F(1, 13) < .01$.

Tests of dishabituation to the broken and complete rods revealed that the infants habituated to the T/NA/NR display did not dishabituate to the broken rod, $t(15) = 1.20$, *ns*, but showed significant dishabituation to the complete rod, $t(15) = 2.63, p < .05$. Again, this pattern of results is inconsistent with perception of object unity. In contrast, the infants habituated to the T/NA/NR control display dishabituated somewhat both to the broken rod, $t(15) = 1.66, p = .12$, and to the complete rod, $t(15) = 2.03, p = .06$.

In sum, these results suggest that misalignment of the rod pieces, even when the rod pieces undergo common motion, disrupts 4-month-olds' perception of object unity. This result generalizes to cases in which the edges of the rod pieces are relatable according to Kellman and Shipley's (1991) criteria. Moreover, there is suggestive evidence that 4-month-olds perceive nonaligned rod pieces, with nonrelatable edges, as disjoint objects.[2] Thus common motion of disparate display elements, along with accretion and deletion of background texture, seem insufficient as information for object unity in some cases.

EXPERIMENT 3

The third experiment examined adults' responses to the T/A/R, NT/A/R, T/NA/R, and T/NA/NR displays. Kellman and Spelke (1983) found that adults' judgments of object unity in rod-and-box displays often differed from the responses of infants. For example, whereas both infants and adults perceived object unity in a display in which the rod pieces underwent common motion relative to a stationary box, only the adults did so when the rod pieces were stationary. Thus it seems likely that infants and adults sometimes rely on different sources of visual information in determining object unity.

Method

Participants. Sixteen adult volunteers served as participants, 13 undergraduate students and 3 graduate students. All were naive to the purpose of the study.

Stimuli, Apparatus, and Procedure. Each student viewed the T/A/R, NT/A/R, T/NA/R, and T/NA/NR displays on the same monitor as used in Experiments 1 and 2, at the same viewing distance. The participants were first shown a pencil and an envelope on a table, arranged such that the two

---

[2]An alternative account of this result might be found by considering the possibility that the amount of dishabituation shown by infants depends to an extent on the amount of new information in test displays, relative to habituation displays. In the case of the T/NA/NR complete rod, there are two new features present, the angles in the rod. Thus, the complete test display in the T/NA/NR condition may appear novel not because it is connected, but rather because it contains new features.

This possibility has been addressed in a new condition for which pilot data have been obtained. In this condition, the rod pieces in the habituation display each contain an angle. The edges at the points of interposition above and below the box are similar in placement and orientation to those in the T/NA/NR display. If infants dishabituate solely on the basis of novel features, they would not be expected to prefer the complete rod after habituation to this new display, because it contains no novel features (moreover, both the complete and the broken rod test displays contain four angles). However, there is a strong preference for the complete rod in this condition, suggesting that the infants experienced the bent rod pieces above and below the box as disjoint objects. This finding provides additional support for our claim that young infants attend to edge relatability in the unit formation process.

ends of the pencil protruded from behind the envelope. They were told that they would be viewing computer-generated displays depicting an arrangement of objects similar to the pencil and envelope. The task was to assign a numerical rating to the strength of the impression of connectedness they received from each display, from 0 for *unambiguously not connected*, to 100 for *unambiguously connected*. Any value between 0 and 100 was permitted, depending on the strength of the impression (e.g., a rating of 50 would indicate an ambiguous percept). All participants agreed that the pencil and envelope arrangement would receive a rating of 100. Order of display presentation followed a balanced Latin-square design.

Results and Discussion

The adult participants' ratings of connectedness in the displays were as follows: T/A/R, 93.4 (13.5); NT/A/R, 94.1 (13.4); T/NA/R, 50.9 (32.3); T/NA/NR, 34.7 (33.0). These judgments differed significantly, $F(3, 36) = 38.14, p < .001$. Post hoc tests revealed that the T/A/R and NT/A/R displays were not judged differently in terms of connectedness, $t(15) = .15$, *ns*. However, the T/NA/R display received a significantly lower rating of connectedness than did the T/A/R display, $t(15) = 4.29, p < .01$. The T/NA/NR display received a lower rating of connectedness than did the T/NA/R display, $t(15) = 1.80, p < .05$ (one-tailed).

These results indicate that the adult participants did not differentiate the T/A/R and NT/A/R displays in terms of the impression of connectedness of the two rod pieces. Thus accretion and deletion of background texture may not serve to aid adults in solving the object unity problem. (Most participants reported that they did not even distinguish between the two displays, until the background was pointed out to them.) However, the adults responded in like manner to the infants in terms of relatability of the rod edges. When the edges were nonaligned but relatable (T/NA/R), the displays appeared ambiguous in terms of unit of the rod pieces. When the edges were nonaligned and nonrelatable (T/NA/NR), the judgments tended to favor nonconnectedness of the rod pieces.

GENERAL DISCUSSION

Experiments 1 and 2 demonstrated that young infants' perception of object unity is not dependent on a single visual cue such as common motion of surfaces (Experiment 3 extends this finding to adults). Rather, it seems likely that object unity is multiply specified, by a variety of cues. These include common motion (relative to a stationary occluder), interposition, three-dimensional depth cues (binocular disparity, self-produced motion parallax, and accommodation and convergence), accretion and deletion of background texture, and alignment and relatability of edges. Figure 7 shows

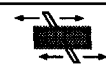
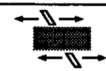
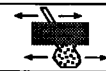
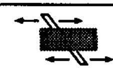
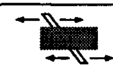
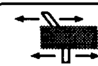
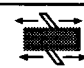
Study	Display	Common (and relative) motion	Interposition	Three-dimensional depth cues	Accretion/deletion of texture	Alignment of edges	Relatability of edges	Resulting Percept
		Visual Information in Display						
Kellman and Spelke (1983) Experiment 1		X	X	X	X	X	X	Single rod
Kellman and Spelke (1983) Experiment 3		X		X	X	X	X	Two rod pieces
Kellman and Spelke (1983) Experiment 5			X	X		X	X	Ambiguous
Kellman and Spelke (1983) Experiment 6		X	X	X	X		X[a]	Single object
Kellman et al. (1986)		X	X	X	X	X	X	Single rod
Slater et al. (1990) Experiment 5		X	X	X		X	X	Single rod
Experiment 1 (T/A/R)		X	X		X	X	X	Single rod
Experiment 1 (NT/A/R and ST/A/R)		X	X			X	X	Ambiguous
Experiment 2 (T/NA/R)		X	X		X		X	Ambiguous
Experiment 2 (T/NA/NR)		X	X		X			Two rod pieces

Figure 7. Schematic depictions of rod-and-box occlusion displays shown to 4-month-old infants in past studies, along with the information in each display for object unity. See text for details.

[a]More precisely, the major axes of the two surfaces in this display were relatable, rather than their edges.

the outcomes of several studies of 4-month-olds' perception of object unity, and the stimulus displays used in each.

To perceive object unity in a rod-and-box display, an observer must extract the depth relations among the display elements (or apparent depth relations, in the case of two-dimensional displays). If such segregation of surfaces is not ascertained, then there might be no need to "fill in" the missing piece of the rod (because it would already be occupied by the box), and perception of object unity might not occur (Slater, 1995). That is, there is information in a two-dimensional display for co-planarity of display elements, from stereopsis, motion parallax, and accommodation and convergence. Although adults clearly perceive the NT/A/R display as a single rod behind a box, it may be that 4-month-olds cannot overcome the cue-conflict inherent in this display (see Johnson & Náñez, 1995).

On the other hand, there is information in the NT/A/R display for at least two depth planes, from interposition and common motion of the rod pieces relative to the box. It may be that information for multiple depth planes was not strong enough (absent the supplementary cue of accretion and deletion of texture) to overcome the information for a single depth plane. Thus surface segregation did not occur, precluding the perception of object unity.

However, 4-month-olds' segregation of surfaces in the object unity task can occur in the absence of accretion and deletion of texture, as demonstrated by Slater et al. (1990). Infants in that study viewed three-dimensional rod-and-box displays against a matte white background, and subsequently looked longer at a broken rod than at a complete rod. It seems likely that the added depth information from stereopsis, self-produced motion parallax, and perhaps accommodation and convergence, supported segregation of surfaces in that display. By 4 months of age, infants are sensitive to small amounts of motion parallax (von Hofsten, Kellman, & Putaansuu, 1992), and many 4-month-olds demonstrate stereopsis (Birch, Gwiazda, & Held, 1982; Fox, Aslin, Shea, & Dumais, 1980). The results of Experiment 1 suggest that limitations in the ability of young infants to segregate surfaces has implications for veridical perception of object properties, such as continuation behind occluders.

To perceive the unity of two disparate surfaces in a partial occlusion display, an observer must determine whether the edges of the two surfaces are relatable (Kellman & Shipley, 1991). However, relatability of edges may not be sufficient to specify connectedness behind an occluding surface to young infants when other cues are missing. Four-month-olds apparently do not perceive object unity in motion displays without three-dimensional depth and background texture (this study, Experiment 1) nor in stationary displays containing these cues (Kellman & Spelke, 1983, Experiment 5; see Figure 7), despite the relatability of edges of the rod pieces in these displays.

Moreover, young infants, like adults, apparently assign nonrelatable edges to disjoint surfaces (this study, Experiment 2).

These findings seem difficult to reconcile with Kellman's (1993) two-process theory of unit formation. There are at least two problems with the theory: the ages at which the primitive and rich processes are functional, and the contributions of each process to unit formation. Consider each problem in turn.

First, it may have been that Kellman (1993) underestimated the age at which the rich process becomes available to young infants in solving the object unity problem. That is, perhaps the rich process contributes to perception of object unity in infants younger than 6 months of age. This would account for the outcome of Experiment 2. However, the results of Experiment 1 cast doubt on this possibility. There was no apparent perception of object unity in the NT/A/R and ST/A/R displays, despite common motion of surfaces and relatability of edges, which should provide support for both the primitive and rich processes, respectively.

These results call into question the second aspect of Kellman's (1993) theory outlined before; that is, the contributions of each process to unit formation. Even when visual information is present in support of both primitive and rich processes, if other cues are absent (e.g., supplementary depth information), unit formation can be disrupted.

In our view, a more useful account of young infants' perception of object unity can be found by considering recent studies of adults' perception of partly occluded surfaces (e.g., Nakayama, Shimojo, & Silverman, 1989). Nakayama, Shimojo and colleagues noted that perceptual completion of one surface behind another depends on two subprocesses, depth placement and contour ownership (Nakayama et al., 1989; Nakayama & Shimojo, 1990). That is, the observer must ascertain relative depth ordering of surfaces in a display, and determine whether the edges of surfaces that appear to lead behind the occluder are likely to be joined behind it. If either subprocess is disrupted, then the partly occluded surface will not be experienced as complete behind the occluder.

This supports our view that the unit formation process in young infants is multiply determined. It may be that a certain threshold of visual information is required for young infants to solve the object unity problem, in terms of depth placement and contour ownership. That is, a *threshold model* would stipulate that insufficiency of cues may often be the best account of an apparent failure to perceive object unity in a particular display. Evidence in favor of the threshold model was recently obtained by Johnson and Aslin (1995) in a study of 2-month-olds' perception of object unity. Unlike 4-month-olds, 2-month-olds appear to perceive the T/A/R display as ambiguous, in terms of the unity of the rod pieces (Johnson & Náñez, 1995, Experiment 2). It may be that the threshold for perception of object unity is lower in older infants. The threshold seems to be even lower in adults (present study, Experiment 3).

Johnson and Aslin (1995) raised the level of visual information for object unity above 2-month-olds' threshold by revealing more of the rod behind the box, relative to the original T/A/R display (i.e., the T/A/R display employed by Johnson & Náñez, 1995; this display was of similar dimensions to the T/A/R display used in Experiment 1 of this study). This was accomplished in two ways: by reducing the box height, and by placing gaps in the box at strategic locations. The 2-month-olds observed by Johnson and Aslin (1995) showed strong evidence of perception of the unity of the rod behind the box. Thus given sufficient visual information, infants as young as 2 months of age appear capable of unit formation in partial occlusion displays.

Several important details of the threshold model remain unclear at present, such as the relative importance of each cue. For example, it would be interesting to know if infants would perceive the unity of two rod parts, whose edges were not relatable, in three-dimensional displays. Perhaps the added information for surface segregation in a 3-D display would support perception of object unity.

Another important question is why the threshold decreases with age. It may be that with increased visual experience, there are improvements in the use of visual cues to disambiguate potentially uncertain object relations. For example, perhaps infants become better able to attend to multiple visual cues simultaneously. Alternatively, perhaps with richer knowledge of objects, infants are able to generalize from object relations experienced in the world to occlusion displays viewed in experimental situations.

Spelke (1990) noted that any mechanism for segmenting the visual array into objects must ascertain the boundaries of adjacent objects, the complete shapes of partly occluded objects, and the continued existence of objects that are no longer visible. The studies shown here broaden our knowledge of limitations and abilities in the ways in which young infants accomplish these tasks. Infants' organization of visual scenes clearly utilizes a variety of the cues available. It remains a goal for future research to elucidate the specific cues used in this process, and the developmental progression of cue use. Methodologies incorporating cue-conflicts, such as the present experiments, hold promise in reaching this goal.

REFERENCES

Birch, E.E., Gwiazda, J., & Held, R. (1982). Stereoacuity development for crossed and uncrossed disparities in human infants. *Vision Research*, *22*, 507–513.

Craton, L.G. (1994, June). *Infants' knowledge of support relations and the perception of partially occluded objects*. Paper presented at the International Conference on Infant Studies, Paris.

Fox, R., Aslin, R.N., Shea, S.L., & Dumais, S.T. (1980). Stereopsis in human infants. *Science, 207*, 323–324.

Hofsten, C. von, Kellman, P.J., & Putaansuu, J. (1992). Young infants' sensitivity to motion parallax. *Infant Behavior and Development, 15*, 245–264.

Horowitz, F.D., Paden, L., Bhana, K., & Self, P. (1972). An infant-control procedure for studying visual fixations. *Developmental Psychology, 7*, 90.

Johnson, M.H. (1990). Cortical maturation and the development of visual attention in early infancy. *Journal of Cognitive Neuroscience, 2*, 81–95.

Johnson, S.P. (1995, May). *Young infants' perception of surfaces in depth: Transparency, texture, and illusory contours.* Poster session presented at the annual Association for Research in Vision and Ophthalmology conference, Ft. Lauderdale, FL.

Johnson, S.P., & Aslin, R.N. (1995). Perception of object unity in 2-month-old infants. *Developmental Psychology, 31*, 739–745.

Johnson, S.P., & Náñez, J.E. (1995). Young infants' perception of object unity in two-dimensional displays. *Infant Behavior and Development, 18*, 133–143.

Kellman, P.J. (1993). Kinematic foundations of visual perception. In C. Granrud (Ed.), *Visual perception and cognition in infancy* (pp. 121–173). Hillsdale, NJ: Erlbaum.

Kellman, P.J., & Shipley, T.F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology, 23*, 141–221.

Kellman, P.J., & Spelke, E.S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology, 15*, 483–524.

Kellman, P.J., Spelke, E.S., & Short, K.R. (1986). Infant perception of object unity from translatory motion in depth and vertical translation. *Child Development, 57*, 72–86.

Koffka, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt, Brace, & World.

Nakayama, K., & Shimojo, S. (1990). Toward a neural understanding of visual surface representation. *Cold Spring Harbor Symposia on Quantitative Biology, 40*, 911–924.

Nakayama, K., Shimojo, S., & Silverman, G.H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception, 18*, 55–68.

Slater, A. (1995). Visual perception and memory at birth. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in infancy research.* (vol. 9, pp. 107–162) Norwood, NJ: Ablex.

Slater, A., Johnson, S.P., Brown, E., & Badenoch, M. (1996). Newborn infants' perception of partly occluded objects. *Infant Behavior and Development, 19*, 147–150.

Slater, A., Morison, V., Somers, M., Mattock, A., Brown, E., & Taylor, D. (1990). Newborn and older infants' perception of partly occluded objects. *Infant Behavior and Development, 13*, 33–49.

Slater, A., Johnson, S.P., Kellman, P.J., & Spelke, E.S. (1994). The role of three-dimensional depth cues in infants' perception of partly occluded objects. *Early Development and Parenting, 3* 187–191.

Spelke, E.S. (1990). Principles of object perception. *Cognitive Science, 14*, 29–56.

Spelke, E.S., Breinlinger, K., Jacobson, K., & Phillips, A. (1993). Gestalt relations and object perception: A developmental study. *Perception, 22*, 1483–1501.