Learning Visual Units After Brief Experience in 10-Month-Old Infants

Amy Needham, a Robert L. Goldstone, b Sarah E. Wiesen a

aDepartment of Psychology and Human Development, Vanderbilt University
bDepartment of Psychological and Brain Sciences, Indiana University

Received 10 April 2013; received in revised form 7 August 2013; accepted 26 August 2013

Abstract

How does perceptual learning take place early in life? Traditionally, researchers have focused on how infants make use of information within displays to organize it, but recently, increasing attention has been paid to the question of how infants perceive objects differently depending upon their recent interactions with the objects. This experiment investigates 10-month-old infants’ use of brief prior experiences with objects to visually organize a display consisting of multiple geometrically shaped three-dimensional blocks created for this study. After a brief exposure to a multipart portion of the display, each infant was shown two test events, one of which preserved the unit the infant had seen and the other of which broke that unit. Overall, infants looked longer at the event that broke the unit they had seen prior to testing than the event that preserved that unit, suggesting that infants made use of the brief prior experience to (a) form a cohesive unit of the multipart portion of the display they saw prior to test and (b) segregate this unit from the rest of the test display. This suggests that infants made inferences about novel parts of the test display based on limited exposure to a subset of the test display. Like adults, infants learn features of the three-dimensional world through their experiences in it.

Keywords: Perceptual learning; Object perception; Unitization; Infant perception

1. Introduction

When we look at the environment surrounding us, we see a collection of objects that make sense to us. Although the light entering our eyes contains a continuous array of different wavelengths and intensities, we perceive a discrete set of objects of various sizes,
shapes, and colors. How are we able to so effortlessly process what we see into distinct objects? One major contributor to this ability is perceptual learning.

Perceptual learning is the process through which the visual system detects regularities in the environment, internalizes these regularities, and uses these internalized patterns to improve its ability to subsequently respond to the same environment. This process can occur in both supervised conditions (i.e., learning that is reinforced through immediate feedback) and unsupervised conditions (i.e., perceptual regularities that are detected under conditions in which feedback is not provided; Goldstone, 1998).

Existing studies provide evidence that adults create visual units on the basis of unsupervised learning. Behrmann, Zemel, and Mozer (1998) found that judgments about the featural similarities between two parts were faster when the two parts belonged to the same object rather than different objects. Further work found an influence of experience on subsequent part comparisons. Two stimulus components were interpreted as belonging to the same object if they co-occurred many times (Zemel, Behrmann, Mozer, & Bavelier, 2002). Although general principles are useful in most situations, there are exceptions to these principles. Object fragments that are not naturally grouped together because they do not follow the Gestalt law of good continuation can nonetheless be perceptually joined if participants are familiarized with an object that unifies the fragments, as is shown in Fig. 1 (see also Pevtzow & Goldstone, 1994).

Previous work on perceptual learning has indicated that adults create functional perceptual units for visual components that frequently co-occur, particularly if their co-occurrence is relevant for a task (Goldstone, 2000). When participants learned a conjunctive search task in which three line segments were needed to distinguish the target from distracters, impressive and prolonged decreases in search slopes were observed over 20 h-long sessions (Shiffrin & Lightfoot, 1997). This is consistent with the hypothesis that conjunctive training leads to unitization of the set of diagnostic line segments. Furthermore, subsequently presented objects tend to be interpreted in terms of these unitized

![Fig. 1. Stimuli used by Zemel et al. (2002). People make efficient judgments when comparing 2 two-humped segments (panel A) that are naturally perceived as belonging to the same occluded object. Fragments that do not follow the Gestalt law of good continuation (panel B) are not naturally perceived as belonging to the same object. Nonetheless, such fragments (panel B) can become perceptually joined if participants are familiarized with a connecting segment that unifies the fragments (panel C) [Permission to reproduce figure granted by author].](image)
components (Austerweil & Griffiths, 2013). Goldstone, Rogosky, Pevtzow, and Blair (2005) report a study in which participants first learned to categorize objects on the basis of particular sets of line segments. Subsequently, they were given a perceptual part-whole judgment task that required them to determine whether a part, either a learned unit or not, was present in a whole pattern. Whole objects were more likely to be broken down into parts that were relevant during categorization.

Because much of the learning that happens early in life is not facilitated by feedback, unsupervised perceptual learning would be an effective way for infants to learn about the visual world (Bhatt & Quinn, 2011; Fiser & Aslin, 2001, 2002; Kirkham, Slemmer, & Johnson, 2002; Saffran, 2001). Learning visual building blocks should be an important task for infants, who are faced with cataloging all of the objects that populate their worlds. Prior studies indicate that 4.5-month-old infants made use of even a brief (5 s) prior experience with an object that subsequently appeared as part of a composite test display (Needham & Baillargeon, 1998). This study showed that 5 s of experience with part of the test display was sufficient in enabling infants to correctly parse an otherwise ambiguous test display. These experiences were useful over a relatively long period of time. Infants were able to utilize just 2 min of prior experience with a box to parse a display containing that box 24 h later (Needham & Baillargeon, 1998).

It is not clear whether infants would use prior experience with a pair of objects to group the objects into a unit that they would then segment from a larger display. This is an important question because in the world outside the lab, segmentation problems do not come at the infant one at a time. To get a better sense of what infants actually experience on a day-to-day basis, we need to understand how infants use their prior experiences to group some object parts together and segment others into separate units, all within the same scene. It is also not clear what kinds of experiences would be necessary for infants to engage in such a process. Perhaps more extensive amounts of experience would be necessary for this more complicated process to take place.

Previous infant research has shown that infants’ exposure to a shape may influence their subsequent interpretation of displays containing that shape (Quinn & Schyns, 2003; Quinn, Schyns, & Goldstone, 2006; see Fig. 2). We built on this work by asking whether infants would form a unit of two differently shaped three-dimensional objects and then parse that unit from a larger set of objects.

In the current study, we examined whether three brief exposures to a subset of a Test Display would lead infants to create expectations about this pair of objects but also about parts of the Test Display they had not yet experienced. In other words, we wanted to know whether infants would (a) unitize the two objects they were briefly exposed to during the familiarization phase of the study and (b) generalize this learning to form globally consistent expectations about the novel and the familiar parts of the Test Display.

The goal of the current study, then, is to see if learning a particular unit leads infants to interpret the rest of their world so as to be consistent with this unit. If their entire visual world at one moment consists of ABCDE, then prior experience with Subset BC is consistent with other units such as Subset DE, but it is inconsistent with Subset AB, which is a
visual parsing that cuts across the previously learned Subset BC. Given evidence suggesting that infants process their visual worlds in a fragmentary or local fashion (Smith, 2009; Stiles, 2008), it is by no means obvious that infants have the capacity or even tendency to create globally consistent parsings or organizations of their visual worlds.

For this study, it was important to use a display that could be seen as one large object or as a collection of smaller objects. To accomplish this goal, we created five three-dimensional objects that each had a different shape and a visible boundary (surface discontinuity) between each object that infants could presumably detect (4-month-old infants did so in Kaufman & Needham, 2010). The five objects could also be seen as a single unit based on proximity, color, and pattern (see Fig. 3). All five objects together comprised the Test Display, which was divided into smaller chunks for use in the familiarization and test phases of the study.

Fig. 3. Test Display showing all five objects in order. During the familiarization phase, infants were familiarized with a subset of this display (either Subset BC or Subset CD) displayed at three locations on three consecutive trials. Following this, all infants were familiarized with the entire Test Display, as pictured above.
Prior to testing, infants saw a familiarization display (two of the five objects comprising the complete Test Display) on three consecutive trials in three different positions on the apparatus floor (see Fig. 3; the letters are to help explain what infants saw, and these letters did not appear on the objects in the study). On each of these trials, the infants saw the same subset of the Test Display: either Subset BC (objects B and C), or Subset CD (objects C and D). The subsets in each of the familiarization displays looked identical to when they later appeared as part of the Test Display. The familiarization displays were situated in a stationary position on the apparatus floor and did not move during these trials. The infants did not see the objects in motion, but we know that even 3-month-old infants regard position changes of objects that are consistent with the objects having moved together or separately as informative about the composition of the display (Kestenbaum, Termine, & Spelke, 1987). So, by 10 months of age, infants should be able to use this change in position to determine that the two objects comprise a unit. As discussed above, frequent co-occurrence is probably a useful cue for adults and infants when creating new building blocks for organizing the environment into objects.

After the familiarization phase, the study procedure was identical for participants across the two familiarization conditions.

After receiving their three familiarization trials with Subset BC or Subset CD, infants completed one test trial viewing the entire stationary Test Display. Next, infants saw two test events (order was counterbalanced) in which a gloved hand entered the apparatus from a hole in either the left or right wall of the apparatus and grasped the object (either object A or object E) on the end of the Test Display closest to the side of the apparatus from which the hand entered. When object A was grasped and pulled to the left (from the infant’s perspective), object B moved along with it, as one rigid unit (Subset AB). Likewise, when object E was grasped and pulled to the right, object D moved along with it, as one rigid unit (Subset DE).

Our predictions were straightforward. If the brief exposure to the familiarization display (with either Subset BC or Subset CD) allowed infants to form a unit that biased their subsequent parsing of the Test Display, they would see the familiar subset as a separate entity within the Test Display and would expect this subset to remain intact. Thus, the infants would show reliably longer looking at the test event that breaks the subset they experienced in the familiarization portion of the procedure than the event that does not involve the portion of the display they had experienced. Specifically, the infants who saw Subset BC in familiarization would look reliably longer toward the test event in which Subset AB moved aside from the rest of the Test Display (breaking the familiar Subset BC) than they would look toward the test event in which Subset DE moved aside from the rest of the Test Display (about which they received no specific information). Likewise, the infants who saw Subset CD in the familiarization phase would look reliably longer toward the test event in which Subset DE moved aside from the rest of the Test Display (breaking the familiar Subset CD) than they looked toward the test event in which Subset AB moved aside from the rest of the Test Display (about which they received no specific information).
An alternate prediction is that infants would look longer at the test event that brought their attention to the portion of the display that was more novel to them because they did not see it in familiarization. According to this hypothesis, infants would show the opposite pattern of results than the one just described. Specifically, the infants who saw Subset BC in the familiarization phase would look reliably longer at the test event in which Subset DE moved apart than at the test event in which Subset AB moved apart. And likewise, the infants who saw Subset CD in the familiarization phase would look reliably longer at the test events in which Subset AB moved apart than at the test event in which Subset DE moved aside from the rest of the Test Display.

2. Method

2.1. Participants

Thirty-two healthy, full-term infants (14 females) ranging in age from 9 months, 2 days to 10 months, 27 days ($M = 9$ months, 29 days) participated in this study. Thirteen infants received Subset BC familiarization and 19 received Subset CD familiarization. Data from an additional 15 infants were excluded from analyses for the following reasons: fussiness ($n = 5$), inattention (i.e., distracted by observers, $n = 5$), experimenter error ($n = 3$), and the infants’ eyes were not visible for coding purposes ($n = 2$). Eight of these infants were in the Subset BC familiarization group, and seven were in the Subset CD familiarization group.

2.2. Apparatus

Infants were tested individually, sitting on a parent’s lap in front of a display box measuring 201 cm tall, 107 cm wide, and 49.5 cm deep. Infants faced an opening 56 cm tall and 94 cm wide in the front wall of the apparatus. Dark blue cardboard covered with a thin piece of transparent Plexiglas lined the floor of the apparatus. The back wall was medium blue. Openings on the right and left sides of this apparatus allowed the experimenter to reach into the display to manipulate objects. Infants were seated approximately 80 cm from the front surface of the apparatus.

Five three-dimensional objects made of foam board comprised the Test Display. Each object was 17.5 cm at its tallest point, 13 cm wide, and 9.5 cm deep. All objects were red with small blue dots. When in its starting position on the apparatus floor, the Test Display stood 22 cm from the front edge of the apparatus; its right and left edges were 20 cm from the right and left side walls of the apparatus, respectively.

2.3. Events and displays

2.3.1. Familiarization displays

Infants saw a series of three trials featuring a subset of objects that would appear adjacent to each other within the entire Test Display. One familiarization display was
composed of Subset BC (objects B and C); the other was composed of Subset CD (objects C and D). The two objects were firmly attached to form the subsets. The familiarization manipulation was between subjects: Infants saw one or the other of these familiarization displays throughout the familiarization period.

Each familiarization display remained motionless for the duration of the trial. Familiarization displays were moved to new locations between trials while the display was outside the infant’s view. The familiarization displays appeared in three separate locations in consecutive familiarization trials: approximately 18 cm from the left wall of the apparatus, approximately in the center of the apparatus, and approximately 18 cm from the right wall of the apparatus (all were 22 cm from the front edge of the apparatus). The order in which the object pairs appeared during these trials was counterbalanced across all infants in the study. These trials lasted until the infant accumulated 5 s of looking at the familiarization display and then looked away for 2 consecutive seconds or until the infant accumulated 30 s of looking at the familiarization display. Infants looked toward the familiarization displays an average of 7.8 s per familiarization trial.

2.3.2. Test display

The first test trial introduced infants to the complete Test Display (ABCDE), which was placed in the center of the apparatus. This trial lasted until the infant accumulated 10 s of looking at the Test Display and then looked away for 2 consecutive seconds or until the infant accumulated 30 s of looking at the Test Display. Infants looked toward the Test Display an average of 13.65 s during this first test trial.

2.3.3. Moves-left event

The two leftmost objects (see Fig. 4) moved 12 cm away from the rest of the Test Display, creating a noticeable gap. Object A was pulled by a gloved hand and object B

Fig. 4. During the test phase, infants saw two test events. In the Moves-left Event, Subset AB was pulled away from the remainder of the Test Display (Subset CDE). In the Moves-right Event, Subset DE was pulled away from the remainder of the Test Display (Subset ABC). Subsets were shown rigidly connected, without any small relative movements.
moved with it as one rigid unit, Subset AB. This motion took 2 s. After a 1-s rest, the hand pushed Subset AB back to its starting position (adjacent to the rest of the Test Display) and rested on the apparatus floor for 2 s. This event repeated until the infant accumulated 8 s of looking at the subset or Test Display and then looked away for 2 consecutive seconds or until the infant accumulated 60 s of looking at the test event.

2.3.4. Moves-right event

The Moves-right Event was identical to the Moves-left Event, with the following exceptions. Instead of grasping object A, the gloved hand grasped object E. Instead of object B moving as one rigid unit with object A, object D moved as a rigid unit along with object E. In other words, in the Moves-right Event, the hand reached in and separated Subset DE from the rest of the Test Display, whereas in the Moves-left Event, the hand reached in and separated Subset AB from the rest of the Test Display.

2.4. Procedure

Each session began with three trials in which a familiarization display (either Subset BC or Subset CD) was seen in three different locations on the otherwise empty apparatus floor. After three trials showing the same familiarization display, all infants received one test trial showing the entire Test Display.

The last two test trials featured events in which subsets of the Test Display were separated from the rest of the Test Display: One test trial featured the Moves-left Event, and the other test trial featured the Moves-right Event. Each of these trials featured one of these events repeating from the beginning until the end of the trial. The order of these two test trials (Moves-left Event first or Moves-right Event first) was counterbalanced across infants.

3. Results

A preliminary analysis of variance (ANOVA) was conducted to determine whether the sex of the infants had a significant effect on their looking times at the test events. This analysis yielded no significant main effects or interactions (ps > 0.05), so we collapsed data across sex for subsequent analyses.

Twenty-two of the 32 infants in the final sample looked longer at the test event that broke the Subset they had seen during familiarization than at the test event that preserved that Subset. This was significant when tested with a sign test, \( p = .025 \).

Infants’ looking times toward the test events were analyzed by means of ANOVA. Familiarization display (Subset BC or Subset CD) and test order (which test event came first: either the Moves-right Event first or Moves-left Event first) were between subjects factors. Test event (whether the test trial featured the Moves-right Event or Moves-left Event) was a within subjects factor. This analysis produced a significant interaction
between test event and familiarization display, $F(1, 28) = 9.43$, $p = .005$, as well as a significant interaction between test event and test order, $F(1, 28) = 5.57$, $p = .025$ (see Fig. 5).

A follow-up analysis of covariance (ANCOVA) was conducted with test order as a covariate, and this analysis produced a significant effect of test event $F(1, 29) = 7.00$, $p = .013$. This analysis also yielded significant interactions between test event and familiarization display, $F(1, 29) = 9.27$, $p = .005$, and between test event and test order, $F(1, 29) = 5.19$, $p = .030$. These findings show that even when controlling for the variability produced by the order of the test events, infants still showed a significant difference in their responses to the test events depending upon which familiarization display they saw prior to testing. Specifically, infants who had seen the familiarization display containing Subset BC looked longer at the Moves-left Event ($M = 44.13$ s) than at the Moves-right Event ($M = 37.43$ s). In contrast, infants who had seen the familiarization display containing Subset CD looked longer at the Moves-right Event ($M = 47.57$ s) than at the Moves-left Event ($M = 30.54$ s).

Infants also had a tendency to look longer at the Moves-right Event ($M = 43.5$ s) than at the Moves-left Event ($M = 36.1$ s). This finding may be related to a general preference for the right side of the body (Baillargeon & Graber, 1988; Michel, 1998; Michel & Harkins, 1986).

Fig. 5. Mean looking times to the Moves-left Event (when Subset AB moved away from the rest of the Test Display) and Moves-right Event (when Subset DE moved away from the rest of the Test Display), depending upon whether the infant had seen the Subset BC or the Subset CD during the familiarization phase of the study. Infants looked more at the test event that broke up the subset they had seen at the beginning of the experiment than they looked at the test event that preserved the subset they were previously exposed to. These findings suggest that even the brief familiarization provided during this study was sufficient to allow infants to form a subset that they expected to be maintained throughout the study, as well as expectations that the other parts of the Test Display would behave in a manner consistent with this learned subset.
4. Discussion

The infants in this experiment showed significantly different patterns of looking at the two test events depending upon their initial experiences viewing familiarization displays containing either Subset BC or Subset CD. When infants saw Subset BC during the familiarization phase, they expected this subset to be maintained over the course of the study and looked longer when this expectation was violated than when it was not. Likewise, when infants saw Subset CD during the familiarization phase, they expected this subset to be maintained throughout the study and looked longer when this expectation was violated than when it was not.

These results demonstrate that infants learn perceptual units through brief experiences with stationary novel objects. Simply seeing a pair of objects together on three occasions (with an average looking time of 7.8 s per trial, or a total of approximately 23.4 s) led infants to group those two objects into a unit that they expected to remain intact. No evidence was found in favor of the novelty-based hypothesis considered in the introduction.

Infants’ familiarization with a subset of the Test Display also led them to form expectations about how novel parts of the Test Display would behave. Specifically, a reasonable prediction of how infants would respond in test is that they would look longer at the event that brought their attention to a novel part of the display (so, the Moves-right Event for the infants who had seen Subset BC in familiarization, and the Moves-left Event for the infants who had seen Subset CD in familiarization). The fact that infants showed the opposite tendency leads us to conclude that, after seeing (for instance) Subset BC in familiarization, they inferred a connection between objects B and C and also inferred a lack of a connection between objects B and A and between objects C and D. Receiving the information that there was no connection between objects C and D was apparently expected, as was the information that there was a connection between objects D and E.

Infants extended their learning during familiarization and formed globally consistent expectations about the novel parts of the Test Display. These findings are unexpected considering that processing speed increases considerably over the course of childhood (Kail, 2007). It is striking that less than 24 s of experience was sufficient for infants to learn a perceptual unit and then use this unit to parse a subsequently viewed display.

Although our results do not allow us to evaluate infants’ versus adults’ relative propensities to learn new visual units, we consider this question a fascinating one and a productive line of inquiry for the future. Our study shows that this ability is quite robust early in development, and we intend to continue asking these bigger-picture developmental questions. Our results indicate that 10-month-old infants are sensitive to the co-occurrence of objects and form units based on their experiences. Furthermore, infants drew conclusions about how novel parts of the Test Display would behave based on their experiences with a subset of the Test Display. This suggests that infants’ ability to form perceptual units is more sophisticated than one might expect, especially considering the large number of unfamiliar stimuli infants encounter on a daily basis.
Simply seeing a pair of stationary objects appear together in three different locations on the apparatus floor was sufficient to prompt infants to infer a connection between the two objects. It seems likely that infants this age could detect the possible separation point between the two objects, due to the shape differences and the surface discontinuity at the boundary seam (these cues are used by 4-month-old infants to assess the likelihood of a separation between two object parts; see Kaufman & Needham, 2010). Because the likelihood of two separate objects appearing together on multiple occasions is small, repeated co-occurrence of an object pair may signal the presence of a unit that should be learned.

Another important component of our findings is they show that infants make use of prior experiences with real, three-dimensional objects in order to parse subsequent displays containing these familiar real objects. Although prior studies suggest that organizational principles and learned features can be used to organize three-dimensional objects (Bhatt & Quinn, 2011), this has not been explicitly demonstrated before now. These findings encourage us to think about infants’ object perception as a fluid process, influenced by brief experiences immediately prior to encounters with objects as well as encounters that occurred 1–3 days prior (Dueker, Modi, & Needham, 2003; Needham & Baillargeon, 1998). Infants are apparently particularly prone to using their previous perceptual experiences as clues to how they should be organizing their subsequent experiences.

5. Conclusions

In the first year of life, infants are faced with many challenges in understanding the visual worlds. Being able to learn units that exist in the world and infer connections among different units would be a major advantage in infants’ attempts to meet these challenges. Their basic visual skills are initially very poor and undergo rapid improvement. These cognitive skills could go a long way toward compensating for these initially poor basic visual capacities. These skills of perceptual learning that require active observation but no concrete feedback from a “teacher” of some kind help infants impose order on their visual environment. This ability to organize their visual surroundings helps infants understand and interact with that world in the most effective way possible.

Acknowledgments

We would like to thank Gabrielle Strouse for her help with the statistical analyses; Klaus Libertus and the undergraduate students in the Infant Perception Lab at Duke University for their help with the data collection; and the parents who generously allowed their infants to participate in this research.
References


No references in this page.


