Cognition 153 (2016) 140–145

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

A smoothness constraint on the development of object recognition

Justin N. Wood

University of Southern California, Department of Psychology, 3620 South McClintock Ave., Los Angeles, CA 90089, United States

ARTICLE INFO

Article history:
Received 19 December 2015
Revised 22 March 2016
Accepted 23 April 2016

Keywords:
Object recognition
Controlled rearing
Newborn
Smoothness
Chicken

ABSTRACT

Understanding how the brain learns to recognize objects is one of the ultimate goals in the cognitive sciences. To date, however, we have not yet characterized the environmental factors that cause object recognition to emerge in the newborn brain. Here, I present the results of a high-throughput controlled-rearing experiment that examined whether the development of object recognition requires experience with temporally smooth visual objects. When newborn chicks (*Gallus gallus*) were raised with virtual objects that moved smoothly over time, the chicks developed accurate color recognition, shape recognition, and color-shape binding abilities. In contrast, when newborn chicks were raised with virtual objects that moved non-smoothly over time, the chicks' object recognition abilities were severely impaired. These results provide evidence for a "smoothness constraint" on newborn object recognition. Experience with temporally smooth objects facilitates the development of object recognition.

1. Introduction

Object recognition is one of the most important functions of the vertebrate visual system. To date, however, the development of object recognition is poorly understood. What environmental factors cause object recognition to emerge in the newborn brain? Does this ability emerge automatically, or do newborns require a specific type of visual input in order to develop accurate object recognition abilities? These types of questions are difficult to address with humans because human infants cannot be raised in strictly controlled environments from birth. In contrast, questions that concern the role of experience in development can be addressed directly with controlled-rearing studies of newborn animals. Here, I describe a high-throughput controlled-rearing experiment that examined whether the development of object recognition requires experience with temporally smooth visual objects.

Researchers have long theorized that biological visual systems leverage the temporal smoothness of natural visual environments to recognize objects (e.g., DiCarlo, Zoccolan, & Rust, 2012; Feldman & Tremoulet, 2006; Foldiak, 1991; Gibson, 1979; Stone, 1996; Wallis & Rolls, 1997; Wiskott & Sejnowski, 2002). In particular, when an object moves smoothly across the visual field, the object projects a series of gradually changing images on the retina. The visual system might take advantage of this natural tendency for temporally contiguous retinal images to belong to the same object by associating patterns of neuronal activity produced by successive retinal images of an object. When provided with temporally smooth visual input, this temporal association process should create object representations that are selective for object identity and tolerant to identity-preserving image transformations (e.g., changes in viewpoint).

A wealth of studies provide evidence that mature visual systems use temporal association mechanisms to create object representations. For example, when human adults are presented with sequential views of an object, the views come to be associated with one another in a manner that aids recognition (Cox, Meier, Oertelt, & DiCarlo, 2005; Liu, 2007; Stone, 1998; Vuong & Tarr, 2004; Wallis, Backus, Langer, Huebner, & Bulthoff, 2009; Wallis & Bülthoff, 2001). Temporal association effects have also been found on the neurophysiological level in adult monkeys (Li & DiCarlo, 2008, 2010; Meyer & Olson, 2011; Miyashita, 1988). In the present study, I examined whether newborn visual systems create more accurate object representations when presented with temporally smooth objects compared to temporally non-smooth objects—as predicted by temporal association models (Wallis, 1998; Wallis & Bülthoff, 2001). Specifically, I examined the *first* visual object representation created by newborn subjects, before their visual systems had been shaped by any prior visual object experience.

1.1. A high-throughput controlled-rearing method

This experiment required controlling all of the subjects' visual experiences from the onset of vision and measuring their object recognition abilities across a range of test trials. To meet these requirements, I used a high-throughput controlled-rearing method

E-mail address: justin.wood@usc.edu

http://dx.doi.org/10.1016/j.cognition.2016.04.013
0010-0277/

(Wood, 2013). The method involves raising newborn chicks in strictly controlled environments and recording their behavior in response to pre-programmed animations (Fig. 1A). We use the term "high-throughput" to describe the method because the controlled-rearing chambers record all of the subjects' behavior (24/7).

I used domestic chicks as an animal model because they are an ideal model system for studying the development of vision (Wood & Wood, 2015a). First, chicks can be raised in strictly controlled environments immediately after hatching, which makes it possible to control all of their visual object experiences. Second, chicks imprint to objects seen in the first days of life. This imprinting behavior can be used to test chicks' object recognition abilities without training (Bateson, 2000; Horn, 2004). Third, birds and mammals process sensory input using homologous neural circuits with similar connectivity patterns (reviewed by Jarvis et al., 2005; Karten, 2013). Since birds and mammals use homologous neural mechanisms to process visual input, controlled-rearing studies of newborn chicks can inform our understanding of the development of both avian and mammalian vision. Finally, chicks develop visual recognition abilities rapidly (Vallortigara, 2012). For example, newborn chicks can begin recognizing objects (Wood, 2013, 2015), faces (Wood & Wood, 2015b), and actions (Goldman & Wood, 2015) at the onset of vision. Newborn chicks can also build integrated object representations with bound color-shape units (Wood, 2014).

In the first week of life (input phase), newborn chicks were raised in environments that contained no objects other than a single virtual object (Fig. 1A). For one group of chicks, the virtual object moved smoothly over time (Temporally Smooth Condition), whereas for another group of chicks, the virtual object moved non-smoothly over time (Temporally Non-Smooth Condition). In the second week of life (test phase), I used an automated two-alternative forced-choice procedure to test the chicks' color recognition, shape recognition, and color-shape binding abilities.

2. Method

2.1. Subjects

Twenty-two domestic chicks of unknown sex were tested. No subjects were excluded from the analyses. The eggs were obtained from a local distributor and incubated in darkness in an OVA-Easy incubator (Brinsea Products Inc., Titusville, FL). After hatching, the chicks were moved from the incubation room to the controlled-rearing chambers in complete darkness. Each chick was raised singly within its own chamber. Ten chicks were raised with a temporally smooth object and 12 chicks were raised with a temporally non-smooth object.[1] This experiment was approved by The University of Southern California Institutional Animal Care and Use Committee.

2.2. Controlled-rearing chambers

The controlled-rearing chambers (66 cm length × 42 cm width × 69 cm height) were constructed from white, high-density polyethylene and were devoid of all real-world (solid, bounded) objects. To present object stimuli to the chicks, virtual objects were projected on two display walls situated on opposite sides of the chamber. The display walls were 19″ liquid crystal display (LCD) monitors (1440 × 900 pixel resolution). Food and water were provided within transparent troughs in the ground (66 cm length × 2.5 cm width × 2.7 cm height). Grain was used as food because it does not behave like an object (i.e., grain does not maintain a rigid, bounded shape). The floors were wire mesh and supported 2.7 cm off the ground by thin, transparent beams. The chambers tracked all of the chicks' behavior (9 samples/s, 24 h/day, 7 days/week) via micro-cameras in the ceilings and automated image-based tracking software (EthoVision XT, Noldus Information Technology, Leesburg, VA). This high-throughput data collection approach allowed us to collect a large number of test trials (168 trials) from each chick, and consequently, measure each subject's object recognition abilities with high precision. In total, 7392 h of video footage (14 days × 24 h/day × 22 subjects) were collected for this experiment.

2.3. Procedure

In the first week of life (input phase), the chicks were raised in controlled-rearing chambers that contained no objects other than a single virtual object (Fig. 1A). The object appeared on one display wall at a time, switching walls every two hours (Fig. 2A). On average, the object measured 9 cm (length) × 7 cm (height) and was displayed on a uniform white background. Half of chicks were imprinted to the object shown in Fig. 1B and half of the chicks were imprinted to the object shown in Fig. 1C.

In the Temporally Smooth Condition (SI Movie 1), the virtual object rotated smoothly around a frontoparallel vertical axis, completing a full rotation every 6 s (30 frames/s). The object had two faces, each with a different color and shape (Fig. 1B). Since the edges of the object (shown during transitions from one face to the other) were identical in color and shape, the object appeared to change smoothly from one 3-D shape to the other 3-D shape. Using this type of geometrically impossible object allowed two different color-shape units to be presented on a single smoothly moving object. Accordingly, I was able to examine whether the first object representation built by newborn chicks contains integrated color-shape units. The same temporally smooth movie was presented throughout the input phase; thus, the transitional probability between images was 1.0.[2]

In the Temporally Non-Smooth Condition (SI Movie 2), the chicks were shown the same virtual object, but the object images were presented in a scrambled order (Fig. 1B). Specifically, I took the 180 unique images (30 frames/s × 6 s) from the temporally smooth animations and randomized the order of the images. On average, the successive images differed by 154° and the minimum difference between two successive views was 50°. To make the images more distinct and eliminate flicker, each image was presented for one second. The same non-smooth movie was presented throughout the input phase; thus, the transitional probability between images was 1.0.

Critically, the virtual objects presented in the two conditions were composed of the same individual images and were equally predictive in terms of the transitional probabilities between images. Furthermore, the subjects received the same amount of overall time with each individual image across the conditions (despite the images being presented at different rates). Thus, any difference in recognition performance between the conditions

[1] The results from the Temporally Smooth Condition were described previously in Wood (2014). In the present study, I directly contrasted chicks raised with temporally smooth objects and temporally non-smooth objects. While the chicks in the two conditions were not tested concurrently, they were tested with the same automated method. Indeed, one major benefit of this controlled-rearing method is that different groups of subjects can be tested in exactly the same way, since the stimuli presentation and data collection processes are fully automated.

[2] The term "transitional probability" refers to the consistency with which the visual images occurred in a particular order. Since the images were presented in a constant order throughput the input phase, the transitional probability between images was 1.0 in both conditions.