



One-shot learning of view-invariant object representations in newborn chicks

Justin N. Wood*, Samantha M.W. Wood

Indiana University, Department of Informatics, 700 N Woodlawn Ave., Bloomington, IN 47408, United States of America



ARTICLE INFO

Keywords:

Object recognition
One-shot learning
Newborn
Chick
Controlled rearing

ABSTRACT

Can newborn brains perform one-shot learning? To address this question, we reared newborn chicks in strictly controlled environments containing a single view of a single object, then tested their object recognition performance across 24 uniformly-spaced viewpoints. We found that chicks can build view-invariant object representations from a single view of an object: a case of one-shot learning in newborn brains. Chicks can also build the same view-invariant object representation from different views of an object, showing that newborn brains converge on common object representations from different sets of sensory inputs. Finally, by rearing chicks with larger numbers of object views, we found that chicks develop enhanced recognition for familiar views. These results illuminate the earliest stages of object recognition, revealing (1) powerful one-shot learning that builds invariant object representations from the first views of an object and (2) view-based learning that enriches object representations, producing enhanced recognition for familiar views.

1. Introduction

To recognize objects successfully, newborns must transform raw retinal inputs into abstract (invariant) object representations that generalize across new viewing situations. While this task feels effortless to human adults, it poses a complex computational challenge because the brain must convert high-dimensional sensory inputs (~1 million photoreceptor cells) into low-dimensional object representations that tolerate real-world image variation (e.g., changes in viewpoint, size, and illumination conditions). How do newborn brains learn to perform this foundational visual task?

One possibility is that newborn brains require an extensive amount of experience with objects before they can begin building invariant object representations. For instance, newborns might need to see objects across a variety of viewpoints to build view-invariant representations, akin to deep convolutional neural networks in artificial intelligence that require thousands to millions of labeled training images to build invariant object representations (Krizhevsky, Sutskever, & Hinton, 2012; Simonyan & Zisserman, 2014; Yamins et al., 2014). An alternative possibility is that newborn brains can build invariant object representations from small amounts of experience. For example, newborns might be capable of “one-shot learning,” building view-invariant representations from a single view of an object. To distinguish between these possibilities, we measured the learning abilities of newborn chicks

during the earliest stages of object recognition—when the chick builds its first object representation.

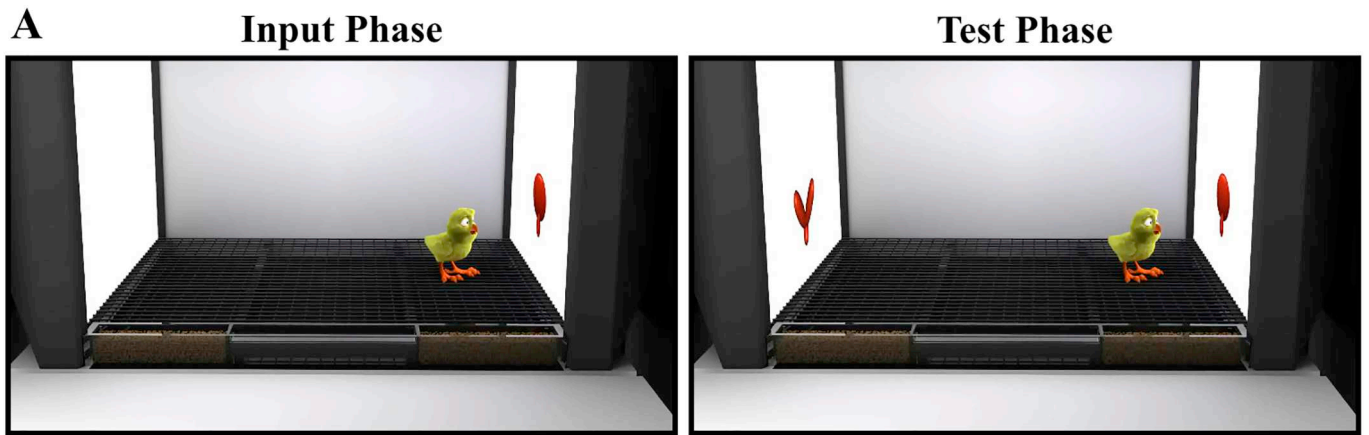
1.1. Using controlled-rearing studies of newborn chicks to study the origins of object recognition

To explore how visual experience shapes object recognition, we used an automated controlled-rearing method with newborn chicks (Wood, 2013). Unlike humans and commonly-used animal models in psychology (e.g., rats, pigeons, monkeys), newborn chicks are precocial, require no parental care, and can be raised in strictly controlled environments from the onset of vision (Fig. 1A). With chicks it is therefore possible to study how specific visual inputs shape the development of object recognition (Wood, 2016; Wood & Wood, 2016). Newborn chicks can also be observed and tested continuously (24/7) for long periods of time, using automated image-based tracking software. Automation allows large amounts of precise behavioral data to be collected from each chick, while eliminating experimenter bias (Wood & Wood, 2019).

Previous automated controlled-rearing studies have shown that newborn chicks develop object recognition rapidly. For instance, chicks can bind color and shape features into integrated object representations (Wood, 2014) and build view-invariant object representations within the first few days of life (Wood, 2013, 2015; Wood & Wood, 2015).

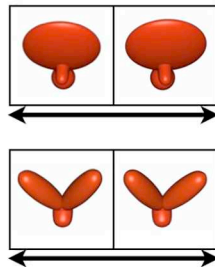
* Corresponding author.

E-mail addresses: woodjn@indiana.edu (J.N. Wood), sw113@indiana.edu (S.M.W. Wood).

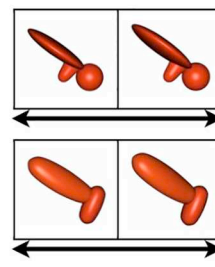


B Viewpoint Ranges Presented During Input Phase

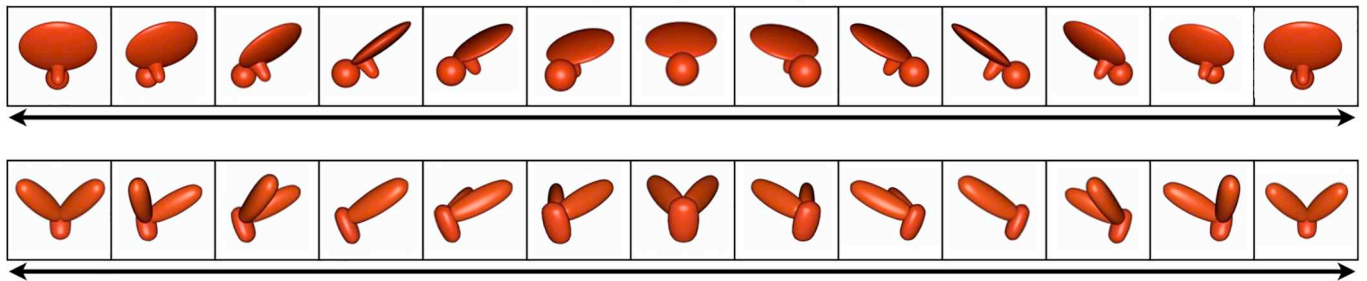
Small Viewpoint Range (front)



Small Viewpoint Range (side)



Large Viewpoint Range



C Test Viewpoints Presented During Test Phase

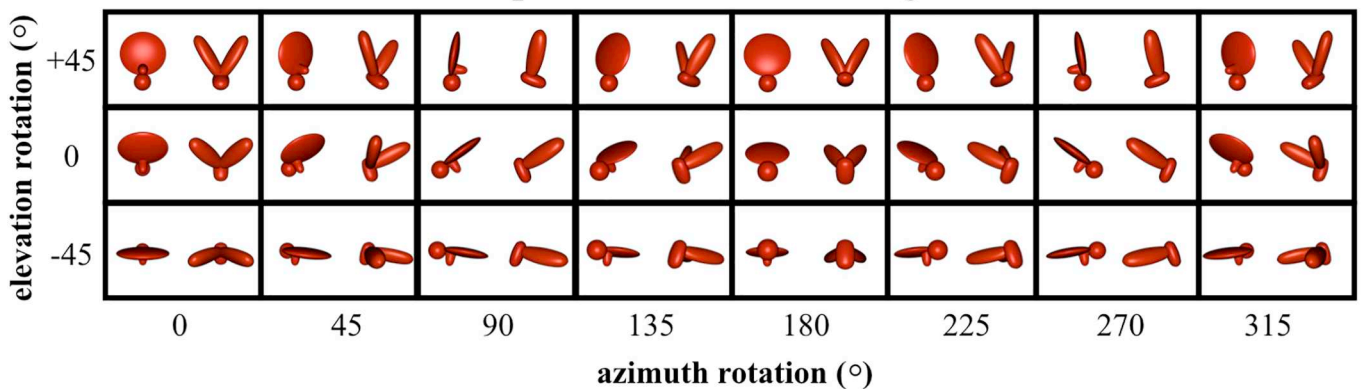


Fig. 1. (A) Illustration of a controlled-rearing chamber during the Input Phase and Test Phase. During the Input Phase, a single virtual object was presented on one display wall. During the Test Phase, the imprinted object was shown on one display wall and an unfamiliar object was shown on the opposite display wall. (B) The object views presented to chicks during the Input Phase. Chicks were reared with an object moving through either a small (11.25°) or large (360°) viewpoint range. The chicks reared with the small viewpoint range saw either a front or side view of the object. (C) The set of viewpoints presented during the Test Phase, consisting of all possible combinations of 3 elevation rotations and 8 azimuth rotations. The imprinted object and unfamiliar object were presented from the same static viewpoint during each test trial to maximize the image-level similarity between the objects.

Newborn chicks can thus solve the complex computational task described above, converting raw retinal inputs into view-invariant object representations. Chicks can build these invariant representations from small amounts of input (e.g., input of a single object rotating through a 60° viewpoint range; Wood, 2013)—an impressive achievement from an artificial intelligence perspective (LeCun, Bengio, & Hinton, 2015).

Although these prior studies indicate that newborn chicks can develop object recognition rapidly from sparse visual inputs, no studies have compared the object recognition abilities of chicks reared with different amounts of visual object input. One hypothesis (hypothesis H1) is that object recognition is initially view-dependent (i.e., newborns can only recognize objects from familiar views) and it becomes gradually more view-invariant as the newborn acquires more experience with object views. According to this hypothesis, visual exploration of shapes is needed to learn to correlate different views of an object into view-invariant representations. This hypothesis predicts that chicks will develop better recognition abilities when reared with larger numbers of object views. An alternative hypothesis (hypothesis H2) is that object recognition is initially view-invariant (i.e., newborns can recognize objects from novel views). According to this hypothesis, view-invariant recognition can emerge from input of a single object view, without the need for learned correlations across views. This hypothesis predicts that chicks will develop view-invariant object recognition when reared with small numbers of object views, and will not necessarily develop better recognition abilities when reared with larger numbers of object views.¹

We tested these predictions by manipulating the amount of visual object input provided to newborn chicks. In the Input Phase (Days 1–5), chicks were reared with a single virtual object rotating through either a small (11.25°) or large (360°) viewpoint range (Fig. 1B). The large viewpoint range displayed 32 times more object views than the small viewpoint range. In the Test Phase (Days 6–14), we then measured the chicks' object recognition abilities by testing their recognition performance across 24 uniformly-spaced viewpoints (Fig. 1C).

To preview our results, we found that chicks successfully created view-invariant object representations when reared with the small viewpoint range, and did not achieve significantly greater levels of view-invariance when reared with the large viewpoint range, consistent with H2. However, we also found that when chicks were reared with the large viewpoint range, they developed enhanced recognition for *familiar* views, indicating that richer visual environments enhance object recognition through view-based learning, consistent with H1. Thus, our results provide evidence for both view-invariant and view-dependent visual learning in newborn brains. In Experiment 2, we extended these results by showing that newborn chicks can build view-invariant object representations from a single view of an object—a case of one-shot learning in newborn animals.

2. Experiment 1

We used a modified version of the controlled-rearing task from Wood (2013). The task involves imprinting chicks to an object moving through a limited viewpoint range, then testing whether the chicks can distinguish their imprinted object from an unfamiliar object across a range of novel viewpoints. We used this task because it meets a reasonable definition of “invariant object recognition” (DiCarlo & Cox, 2007; Zoccolan, Oertelt, DiCarlo, & Cox, 2009). Specifically, chicks can

¹ While prior research demonstrates that newborn chicks can recognize objects across novel views (Wood, 2013, 2015; Wood & Wood, 2015), such data cannot help resolve this debate because the chicks were shown multiple views of the object during imprinting, and therefore had the opportunity to learn to correlate different views into view-invariant representations. By rearing chicks with a small (11.25°) viewpoint range (Experiment 1) and a single view of an object (Experiment 2), we could test whether view-invariant representations emerge only when newborn animals see objects from multiple viewpoints.

recognize the objects in this task across substantial variation in object view (both azimuth and elevation rotations; Wood, 2013). Changing the view of an object produces large and complex changes in the pixel-level appearance of the object. Thus, recognizing objects across novel views requires an invariant representation that can generalize across large, novel, and complex changes in the object's appearance on the retina.

In prior work, we established that chicks' recognition performance in this task cannot be explained by low-level visual representations. Chicks do not use overall brightness as a low-level cue to distinguish between these objects (Wood, 2015), and chicks' recognition performance cannot be accounted for by retina-like (pixel-level) or V1-level representations (Wood, 2013; Wood & Wood, 2015). Moreover, chicks can recognize these objects even when the within-object image difference (i.e., the image-level difference between the test views of the imprinted object and the input views of the imprinted object) is greater than the between-object image difference (i.e., the image-level difference between the test views of the unfamiliar object and the input views of the imprinted object; Wood, 2013; Wood & Wood, 2015)—the hallmark of invariant object recognition (Zoccolan et al., 2009). This task therefore provides a promising starting point for exploring how greater amounts of visual object input shape object recognition.

2.1. Method

2.1.1. Subjects

Thirty-six Rhode Island Red chicks of unknown sex were tested. The sample size ($n = 12$ per group) was determined before the experiment was conducted, based on previous automated controlled-rearing experiments with newborn chicks (Wood, 2013, 2014). No subjects were excluded from the analyses. The eggs were obtained from a local distributor and incubated in an OVA-Easy incubator (Brinsea Products Inc., Titusville, FL). The eggs were incubated in darkness to ensure that no visual input would reach the chicks through their shells. After hatching, the chicks were moved from the incubation room to the controlled-rearing chambers in darkness with the aid of night vision goggles. Each chick was reared singly within its own chamber. This research was approved by The University of Southern California Institutional Animal Care and Use Committee.

2.1.2. Controlled-rearing chambers

The chambers measured 66 cm (length) × 42 cm (width) × 69 cm (height) and contained no real-world (solid, movable) objects (Fig. 1A). To present object stimuli to the chicks, we projected virtual objects on two display walls (19" LCD monitors, with 1440 × 900 pixel resolution) situated on opposite sides of the chamber. Food and water were available within transparent troughs recessed into the ground that measured 66 cm (length) × 2.5 cm (width) × 2.7 cm (height). We used grain as food because a heap of grain does not behave like an object (i.e., a heap of grain does not maintain a rigid, bounded shape). The floors of the chambers consisted of black wire mesh supported over a black surface.

The chicks' location was measured continuously (9 samples/s, 24 h/day, 7 days/week) using micro-cameras in the ceilings of the chambers and automated image-based tracking software (EthoVision XT, Noldus Information Technology, Leesburg, VA). This automated data collection approach allowed hundreds of test trials to be collected from each chick. Consequently, we could measure each chick's object recognition performance with high precision.

2.1.3. Procedure

During the Input Phase (Days 1–5), newborn chicks were reared with a single virtual object (Fig. 1A). We varied the viewpoint range of the imprinted object across chicks. For one group of chicks, the object rotated through a small (11.25°) viewpoint range, whereas for another group of chicks, the object rotated through a large (360°) viewpoint range (Fig. 1B). Thus, the chicks reared with the large viewpoint range

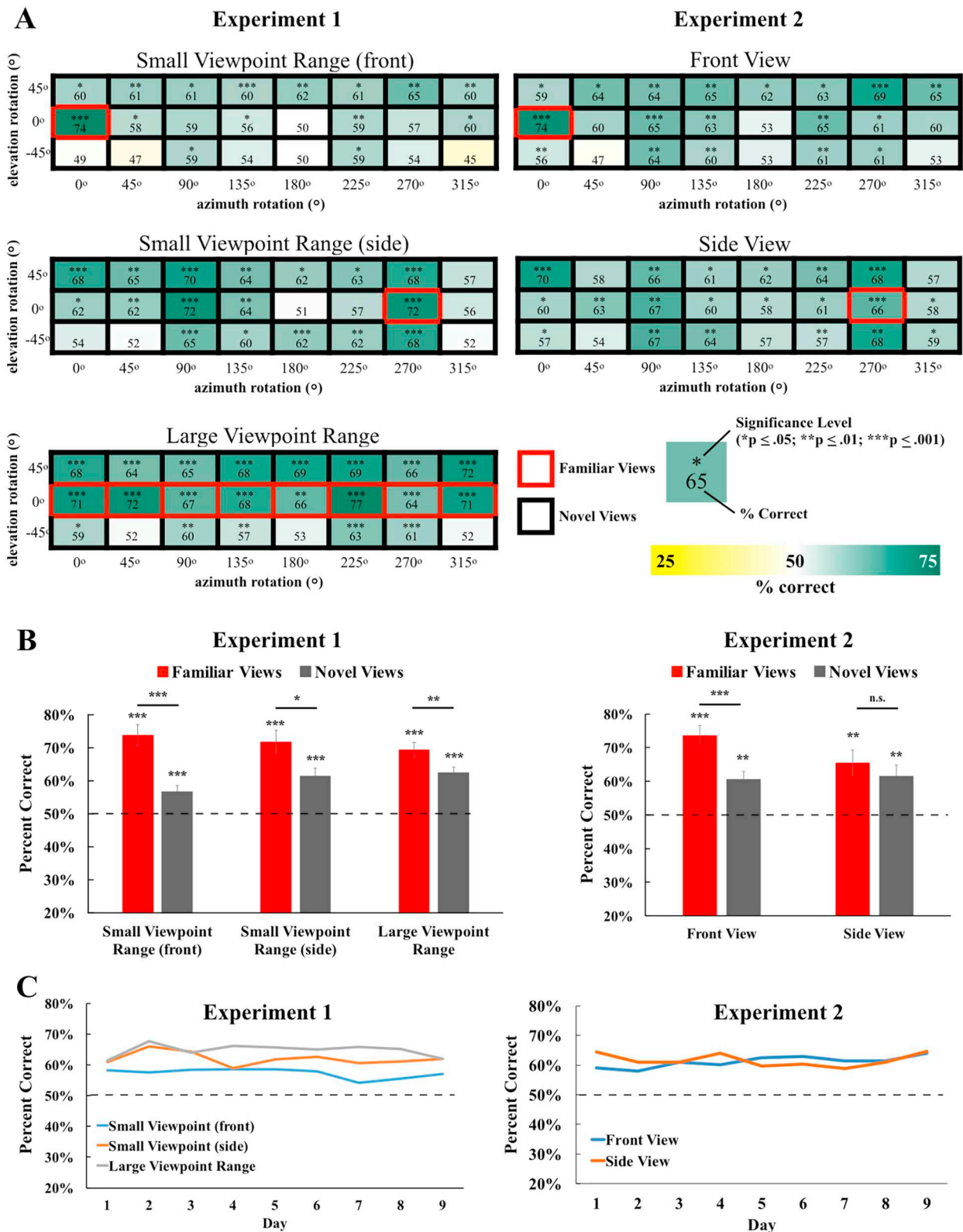


Fig. 2. Results. (A) The chicks' object recognition performance across the 24 test viewpoints in Experiment 1 (left) and Experiment 2 (right). The red boxes denote familiar views and the black boxes denote novel views. Performance (percent correct) is both color-coded and reported as a numeric value, together with its significance according to a one-tailed *t*-test (see key for significance values). (B) The chicks successfully recognized their imprinted object across both familiar views (red bars) and novel views (gray bars). The chicks also showed enhanced recognition when the object was presented from familiar views in both Experiment 1 (left) and Experiment 2 (right). The familiar and novel views used in this analysis are denoted by red and black boxes in Panel A, respectively. Error bars indicate standard error. The dashed lines show chance performance. (C) Overall recognition performance by test day. Performance did not improve over the course of the test phase in Experiment 1 (left) or Experiment 2 (right). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

received 32 times more unique object views than the chicks reared with the small viewpoint range. In addition, to test whether newborn chicks can build the same object representation from different views of an object, the chicks reared with the small viewpoint range were shown either a front view or side view of the object (Fig. 1B). SI Videos 1, 2, 3 show sample animations presented during the Input Phase.

The objects moved at the same speed in the small and large viewpoint range conditions. Since the objects switched direction at the end of their trajectory, the objects in the small viewpoint range condition changed directions at a faster rate than the objects in the large viewpoint range condition. The chicks also saw each object view a greater number of times in the small versus large viewpoint range conditions (since the large viewpoint range contained a larger number of object views, each view was seen a smaller number of times).

On average, the objects measured 8 cm (length) \times 7 cm (height) and were suspended 3 cm off the floor. The object moved at a constant speed (18° of rotation/s). The object was displayed on a uniform white background and appeared on one display wall at a time, switching display walls every 2 h. Half of the chicks were imprinted to Object A, with Object B serving as the unfamiliar object, and half of the chicks were imprinted to Object B, with Object A serving as the unfamiliar object.

During the Test Phase (Days 6–14), we measured the chicks' object recognition performance by testing their recognition abilities across 24 uniformly-spaced viewpoints (Fig. 1C). On each test trial, the imprinted object was shown on one display wall and an unfamiliar object was shown on the other display wall (Fig. 1A). The unfamiliar object was the same size and color as the imprinted object. We measured the amount of time the chicks spent in proximity to each object, by calculating the time spent in zones (22 cm \times 42 cm) next to each object. If the chicks successfully recognized their imprinted object, then they should have spent a greater proportion of time in proximity to the imprinted object than the unfamiliar object.

The objects were shown from 24 test viewpoints, consisting of all possible combinations of 8 azimuth rotations (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°) and 3 elevation rotations (–45°, 0°, +45°). Since the test viewpoints were static images, we flashed the same image every 4 s to draw the chicks' attention to the stimuli. Each viewpoint was presented for 3750 ms, followed by a 250-ms blank display, which repeated for the duration of the trial. This presentation rate is well within newborn chicks' object recognition abilities (i.e., chicks can recognize objects at presentation rates as low as 125 ms per image; Wood & Wood, 2017).

The imprinted object and unfamiliar object were presented from the same viewpoint during each test trial. Thus, the chicks reared with the small and large viewpoint ranges were shown exactly the same stimuli during the test trials, allowing for direct comparisons across the groups. In addition, presenting the two objects from the same viewpoint made the two test images equally novel to the chicks, while also minimizing the image-level difference between the two test images. SI Video 4 shows sample animations presented during the Test Phase.

The test trials lasted 40 min and were separated from one another by 18-min rest periods. During the rest periods, we projected the animation from the Input Phase on one display wall and a white screen on the other display wall. We presented rest periods to measure the strength of the imprinting response and assess ceiling-level performance. Since our task relied on imprinting, ceiling performance was directly constrained by the strength of the imprinting response. The test trials and rest periods were separated by 1-min periods of darkness.

During each day of the Test Phase, the chicks were presented with each viewpoint for one test trial, for a total of 24 test trials per day. Accordingly, each chick received 216 test trials over the course of the experiment. The 24 viewpoints were presented in a randomized order on each test day.

2.2. Results

To measure the chicks' performance, we computed the percent of time each chick spent with the imprinted object versus the unfamiliar object (percent time = time spent with imprinted object / [time spent with imprinted object + time spent with unfamiliar object]). Performance could range from 0% (complete preference for the unfamiliar object) to 100% (complete preference for the imprinted object). Fig. 2A shows the chicks' recognition performance for each individual viewpoint, Fig. 2B shows overall recognition performance across the familiar and novel viewpoints, and Fig. 2C shows recognition performance across the nine test days.

We first examined whether the chicks reared with the large viewpoint range built view-invariant object representations. Pooling across the 24 test viewpoints, the chicks performed significantly higher than chance (50%) level ($M = 65\%$, $SE = 1.7\%$, $t(11) = 8.86$, $p < 10^{-5}$, Cohen's $d = 2.56$). The chicks also performed significantly higher than chance level when the familiar test viewpoints (i.e., the 0° elevation rotation viewpoints; denoted with red boxes in Fig. 2A) were removed from the analysis ($M = 62\%$, $SE = 1.6\%$, $t(11) = 7.55$, $p < 10^{-4}$, Cohen's $d = 2.18$). Thus, the chicks built view-invariant object representations after receiving a large (360°) amount of object rotation experience.

Next, we examined recognition performance for the chicks reared with the small (11.25°) viewpoint range. Pooling across the 24 test viewpoints, the chicks performed significantly higher than chance level when reared with both the front viewpoint range ($M = 57\%$, $SE = 1.7\%$, $t(11) = 4.26$, $p = .001$, Cohen's $d = 1.23$) and the side viewpoint range ($M = 62\%$, $SE = 2.2\%$, $t(11) = 5.37$, $p = .0002$, Cohen's $d = 1.55$). The chicks also performed significantly higher than chance level when the familiar test viewpoints (denoted with red boxes in Fig. 2A) were removed from the analysis, both for the front viewpoint range ($M = 57\%$, $SE = 1.8\%$, $t(11) = 3.79$, $p = .003$, Cohen's $d = 1.09$) and the side viewpoint range ($M = 61\%$, $SE = 2.3\%$, $t(11) = 4.98$, $p = .0004$, Cohen's $d = 1.44$). Despite receiving a small amount of visual object input (11.25° of object rotation experience), the chicks successfully created view-invariant object representations that generalized across novel viewpoints.

We then examined whether performance differed across the chicks reared with the small and large viewpoint ranges. When the analysis included all 24 test viewpoints, the chicks reared with the large viewpoint range performed significantly better than the chicks reared with the small viewpoint range (mean difference = 5%, $t(34) = 2.14$, $p = .04$, Cohen's $d = 0.78$). This effect disappeared, however, when the familiar test viewpoints were removed from the analysis (i.e., the analysis included all of the novel viewpoints received by each group; mean difference = 3%, $t(34) = 1.36$, $p = .18$, Cohen's $d = 0.50$). We observed the same pattern when the analysis only included the same set of novel viewpoints across the groups (i.e., all of the 0° elevation viewpoints were removed from the analysis for all of the groups; mean difference = 3%, $t(34) = 1.36$, $p = .18$, Cohen's $d = 0.50$). Thus, there was no evidence that the chicks achieved greater levels of view-invariance when reared with the large (360°) versus the small (11.25°) viewpoint range. The enhanced recognition performance for the chicks reared with the large viewpoint range can be largely explained by view-dependent learning.

To test for the presence of view-dependent learning more directly, we compared the chicks' recognition performance for familiar and novel test viewpoints (Fig. 2B). The chicks reared with the small viewpoint range performed significantly better when tested with familiar viewpoints versus novel viewpoints, both for the front viewpoint range (mean difference = 17%, $SE = 3.3\%$, $t(11) = 5.10$, $p = .0003$, Cohen's $d = 1.47$) and side viewpoint range (mean difference = 10%, $SE = 4.2\%$, $t(11) = 2.47$, $p = .03$, Cohen's $d = 0.71$). Likewise, the chicks reared with the large viewpoint range performed significantly better when tested with familiar viewpoints versus novel viewpoints

(mean difference = 7%, SE = 1.6%, $t(11) = 4.30$, $p = .001$, Cohen's $d = 1.24$).

Imprinting in chicks is subject to a critical period, which ends approximately 3 days after hatching. Thus, the chicks' representation of their imprinted object was not expected to change over the course of the test phase (which began 5 days after hatching). To test this assumption, we analyzed the proportion of time subjects spent in proximity to their imprinted object as a function of test day. For all conditions, performance was high and significantly above chance level even for the first presentation of the test views (Day 1: two-tailed t -tests, all P s < 0.03), and remained stable over the course of the test phase (Fig. 2C) with little variation as a function of presentation number (repeated-measures ANOVAs with main effect of Test Day, all P s > 0.1). The chicks' recognition behavior was spontaneous and robust, and cannot be explained by learning taking place across the test phase.

Finally, we examined whether the chicks imprinted to the objects at the same rate across the conditions (Fig. 3). We performed a repeated-measures ANOVA with main effects of Imprinting Day and Condition. As expected, we found a main effect of Imprinting Day ($F(4, 84) = 5.066$, $p = .004$), indicating that the imprinting response became stronger across the Input Phase. The main effect of Condition ($F(2, 21) = 0.475$, $p = .628$) and the interaction between Imprinting Day and Condition ($F(8, 84) = 0.443$, $p = .834$) were not significant. Thus, there was no evidence that the chicks imprinted to the objects more strongly, or at a faster rate, when reared with the large (360°) versus the small (11.25°) viewpoint range. This finding indicates that the enhanced recognition performance for the chicks reared with the large viewpoint range cannot be attributed to the amount of time spent with the objects during the Input Phase.

2.3. Discussion

In Experiment 1, newborn chicks developed view-invariant object representations regardless of whether they were reared with a small (11.25°) or large (360°) viewpoint range. The large viewpoint range presented 32 times more object views, but the chicks in the two groups obtained similar levels of view-invariance. Both groups of chicks showed enhanced recognition for familiar views, leading the chicks who were reared with the large viewpoint range to perform better overall (across all of the viewpoints). Experience with greater numbers of object views enhanced object recognition performance, largely through view-dependent learning.

The fact that chicks successfully built view-invariant representations from as little as 11.25° of object rotation experience raises the

possibility that newborn brains are capable of “one-shot learning” of invariant object recognition. To test this possibility, we reared a new group of chicks with just a single view of their imprinted object, then measured the amount of view-invariance achieved by each chick by testing their recognition performance across the same 24 viewpoints used in Experiment 1.

3. Experiment 2

To test whether newborn chicks can perform one-shot learning in a view-invariant object recognition task, we repeated Experiment 1 with one crucial change: rather than presenting an object moving through a small or large viewpoint range, we reared the chicks with a single view of their imprinted object. Since the chicks were imprinted to a single view, there was no opportunity to learn to correlate different views of the object into a view-invariant representation. As in Experiment 1, one group of chicks was reared with a front view of their object, while another group was reared with a side view.

3.1. Method

The methods were identical to those used in Experiment 1, except in the following ways. First, we tested a new group of 20 chicks. No subjects were excluded from the analyses. Half of the chicks were reared with a front view of the object, and half of the chicks were reared with a side view (Fig. 4A). Second, rather than presenting a rotating object in the Input Phase, we presented a single view of the object. To draw the chicks' attention to the object, we flashed the image repeatedly. The image was presented for 3750 ms, followed by a 250-ms blank display between each image (akin to the presentation of the test views in the Test Phase). SI Videos 5, 6 show sample animations presented during the Input Phase.

Additionally, we quantified the image variation across the imprinted views to assess which views were more similar to one another from retina-like (pixel-level) and V1-level perspectives (Fig. 4B). Critically, we selected the imprinted views to ensure that the images of the different objects seen from the same view were more similar to one another than the images of the same object seen from different views. In other words, the *within-object* image distance (i.e., the image-level distance between the two views of the same object) was greater than the *between-object* image distance (i.e., the image-level distance between the same views of the different objects). If newborn chicks build low-level (retina-like or V1-like) representations, then the chicks reared with the same views of different objects should build more similar

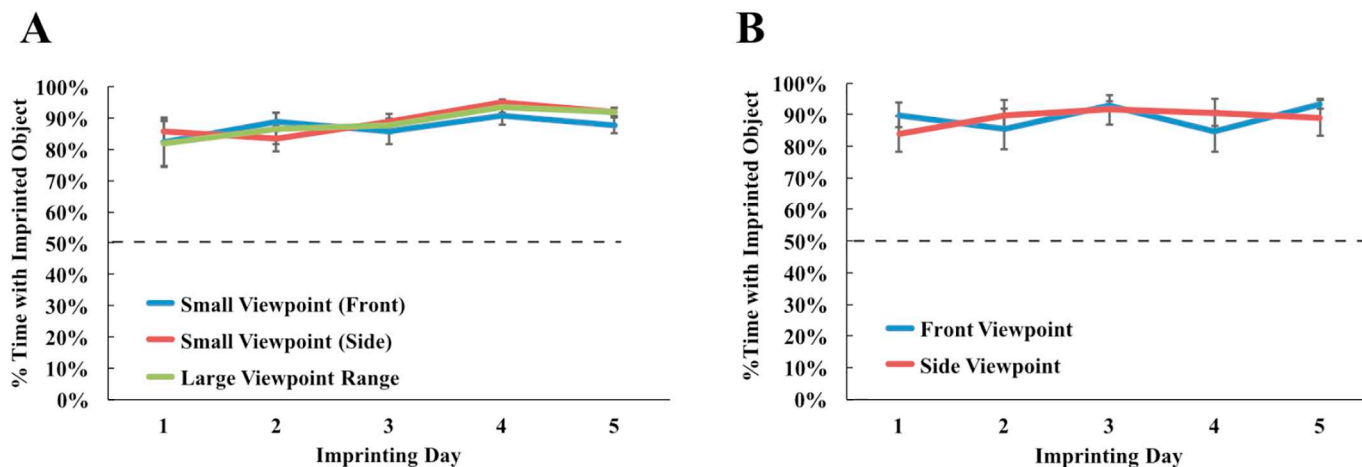


Fig. 3. Imprinting Results. The graphs show the percent of time the chicks spent with their imprinted object across the five imprinting days in (A) Experiment 1 and (B) Experiment 2. There was no evidence that the chicks imprinted to the objects more strongly, or at a faster rate, when reared with the large (360°) versus the small (11.25°) viewpoint range.

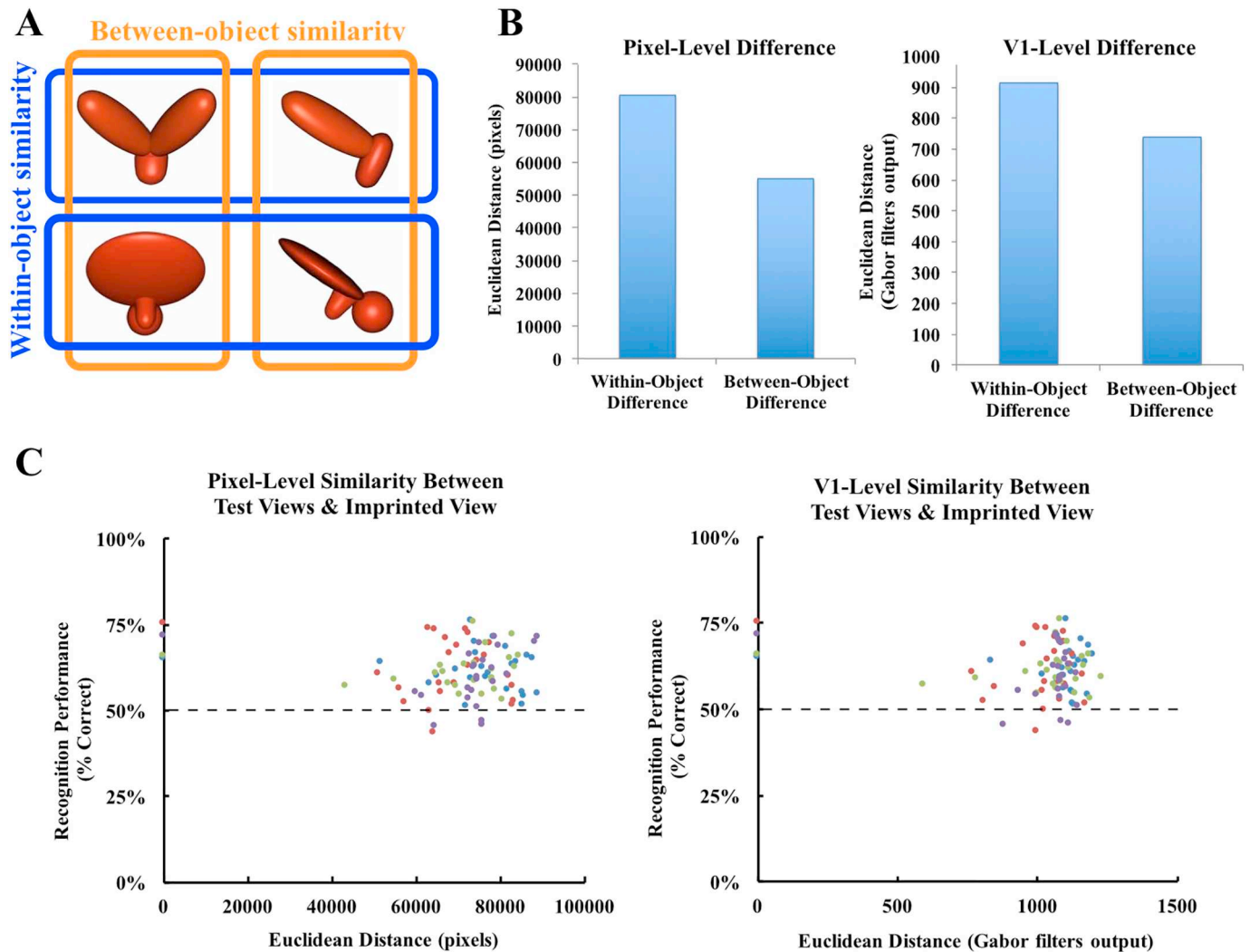


Fig. 4. Image-Level Similarity Analysis. (A) The imprinted views used in Experiment 2. The blue boxes show the images used to compute the within-object image difference (i.e., the image-level difference between the two views of the same object) and the orange boxes show the images used to compute the between-object image difference (i.e., the image-level difference between the two views of the different objects). (B) From both pixel-level and V1-level perspectives, the within-object image difference was greater than the between-object image difference. Thus, if the chicks had relied solely on image-level similarity to recognize objects, then the chicks reared with the same views of the different objects should have built more similar object representations than the chicks reared with the different views of the same object. (C) Recognition performance as a function of the pixel-level and V1-level similarity between the test views and the imprinted view. The dot colors denote the four imprinted images shown in Panel A. Recognition performance was not correlated with either pixel-level or V1-level image similarity, indicating that the chicks developed high-level (view-invariant) object representations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

representations than the chicks reared with different views of the same object. Conversely, if newborn chicks build high-level (view-invariant) representations, then the chicks reared with different views of the same object should build more similar representations than the chicks reared with the same views of different objects.

To measure the image variation across the imprinted views, we computed the within-object image difference and the between-object image difference (Fig. 4A). We quantified image variation in two ways. First, we computed the amount of image variation from a retina-like (pixel-level) perspective. For each image, we (1) measured the brightness level of each pixel in each image, (2) compared the images (i.e., at each pixel location we subtracted the brightness level in one image from the brightness level in the other image), and (3) calculated the Euclidean distance between the images (by squaring the brightness difference at each pixel location and computing the square root of the sum of the squares). As shown in Fig. 4B, the within-object image difference was *greater* than the between-object image difference, showing that the same views of the different objects were more similar (from a

pixel-level perspective) than the different views of the same object.

Second, we computed the amount of image variation from a V1-level perspective. We used a Gabor measure of similarity with the Gabor jet model: a multi-scale, multi-orientation model of V1 complex-cell filtering developed by Lades et al. (1993).² The general parameters and implementation followed those used by Xu and Biederman (2010), which can be downloaded at http://geon.usc.edu/GWTgrid_simple.m. For each image, we measured the magnitude of activation values that the image produced in a set of 40 Gabor jets (8 orientations × 5 scales) at 100 locations across the image. Then, we calculated the Euclidean

² While the Gabor jet model is based on the mammalian primary visual cortex, the avian visual wulst has circuitry and physiological properties that are similar to the mammalian primary visual cortex, and they are widely considered to be homologous structures (Karten, 1969, 2013). For example, like the cat and monkey visual cortex, the visual wulst includes precise retinotopic organization and selectivity for orientation (e.g., Liu & Pettigrew, 2003; Pettigrew & Konishi, 1976).

distance between the vectors to compute the V1-level distance across images. As shown in Fig. 4B, the within-object image difference was greater than the between-object image difference, showing that the same views of the different objects were more similar (from a V1-level perspective) than the different views of the same object. Thus, to succeed, the chicks needed to build invariant object representations that generalized beyond the image-level characteristics of the objects.

3.2. Results

The results are shown in Fig. 2. We first examined whether the chicks successfully built view-invariant object representations. Pooling across the 24 test viewpoints, the chicks performed significantly higher than chance level when reared with both the front viewpoint ($M = 61\%$, $SE = 2.3\%$, $t(9) = 4.92$, $p = .0008$, Cohen's $d = 1.56$) and the side viewpoint ($M = 62\%$, $SE = 3.2\%$, $t(9) = 3.68$, $p = .005$, Cohen's $d = 1.16$). The chicks also performed significantly higher than chance level when the familiar test viewpoint (denoted by red boxes in Fig. 2A) was removed from the analysis, both for the front viewpoint ($M = 61\%$, $SE = 2.3\%$, $t(9) = 4.66$, $p = .001$, Cohen's $d = 1.47$) and the side viewpoint ($M = 62\%$, $SE = 3.2\%$, $t(9) = 3.60$, $p = .006$, Cohen's $d = 1.14$). Despite being reared with just a single view of an object, the chicks successfully developed view-invariant representations that generalized across novel viewpoints.

To test for the presence of view-dependent learning, we compared recognition performance across the familiar and novel test viewpoints. The chicks performed significantly better on the familiar viewpoint than the novel viewpoints when reared with the front viewpoint (mean difference = 13%, $SE = 2.4\%$, $t(9) = 5.43$, $p = .0004$, Cohen's $d = 1.72$), but not when reared with the side viewpoint (mean difference = 4%, $SE = 2.9\%$, $t(9) = 1.40$, $p = .20$, Cohen's $d = 0.44$).

3.2.1. Analysis of image-level similarity between imprinted views and test views

To what extent did recognition performance depend on the image-level similarity between the imprinted views and test views? In prior work, we found that chicks' recognition performance with these two objects does not vary as a function of the pixel-level and V1-level similarity between the imprinted views and test views (Wood, 2013; Wood & Wood, 2015). To explore whether this was also the case in the present experiment—when chicks were imprinted to a single view—we examined whether image-level similarity was predictive of performance (Fig. 4C). To compute image-level similarity, we measured the Euclidean distance between each imprinted view and each test view from pixel-level and V1-level perspectives. If recognition performance was driven by image-level similarity, then we should have observed a correlation between performance and the image-level similarity between the test views and imprinted view.

We first examined whether performance varied as a function of the image-level similarity between the test views of the imprinted object and the imprinted view of the imprinted object (Fig. 4C). Performance was not significantly correlated with pixel-level similarity ($r = -0.128$, $p = .214$) or V1-level similarity ($r = -0.163$, $p = .112$). We observed the same pattern when the familiar test views were removed from the analyses (pixel-level similarity: $r = 0.146$, $p = .165$; V1-level similarity: $r = 0.118$, $p = .262$).

Since we used a two-alternative forced-choice task in which the chicks were always presented with two test images (one image of the imprinted object and one image of the unfamiliar object), we next examined whether performance varied as a function of the image-level similarity between the two test images and the imprinted image. For each test trial, we calculated the image-level similarity between each of the two test images and the imprinted image, and then computed the difference between those values. If performance varied as a function of which test image was more similar to the imprinted image, then performance should have been higher when the test image of the imprinted

object was more similar to the imprinted image than the test image of the unfamiliar object was to the imprinted image. Performance was not significantly correlated with pixel-level similarity ($r = 0.185$, $p = .071$) or V1-level similarity ($r = 0.157$, $p = .127$). We observed the same pattern when the familiar test views were removed from the analyses (pixel-level similarity: $r = -0.069$, $p = .514$; V1-level similarity: $r = -0.145$, $p = .168$). Thus, there was no evidence that the chicks' view-invariant recognition performance can be explained in terms of the image-level similarity between the imprinted views and test views.

3.3. Discussion

Experiment 2 extends the results from Experiment 1 by demonstrating that newborn chicks are capable of one-shot learning in a view-invariant object recognition task. The chicks successfully developed view-invariant representations from a single object view, indicating that newborn brains can generate invariant object representations without the need for learned correlations across views.

4. High-dimensional analysis of newborn object recognition

Can newborn brains build the same invariant object representation from different views of an object? One important feature of our design is that we could directly compare the object representations built by chicks reared with different object views. We could therefore explore whether newborn brains converge on common object representations from different sets of sensory inputs. Our study provided a particularly challenging test case because the within-object image difference was greater than the between-object image difference across the imprinted views (Fig. 4B).

We first compared the overall recognition performance of the two groups. When the analysis included all 24 test viewpoints, performance did not differ significantly when the chicks were reared with the front view versus side view, both in Experiment 1 (mean difference = 4.55%, $SE = 2.82\%$, $t(22) = 1.61$, $p = .12$, Cohen's $d = 0.66$) and Experiment 2 (mean difference = 0.59%, $SE = 3.90\%$, $t(18) = 0.15$, $p = .88$, Cohen's $d = 0.07$). There was no evidence that the chicks developed different object representations when reared with different views of an object.

We next compared the chicks' recognition performance in a high-dimensional recognition space. Each chick's performance across the 24 viewpoints can be thought of as a point in a 24-dimensional space, with each dimension reflecting the chick's performance on one test viewpoint. Comparing performance within a high-dimensional recognition space is informative because a coordinate within a high-dimensional space contains information about the subject's *specific pattern* of successes and failures across the test viewpoints. Conversely, measuring performance within a one-dimensional measurement space (e.g., overall percent correct) does not.³

To calculate the position of each chick's performance within the 24-dimensional recognition space, we computed the percent of time the chick spent with Object A versus Object B for each test viewpoint (dimension). Thus, unlike the results described above, this high-dimensional analysis measures behavior in terms of the chick's preference for Object A (rather than in terms of the chick's preference for their imprinted object). These values were then scaled to fit on the 24 axes of the recognition space, such that 0 indicated chance performance, 1.0 indicated a complete preference for Object A, and -1.0 indicated a complete preference for Object B.

³For example, it is possible to obtain 75% correct performance in a variety of ways (e.g., any combination of 18 correct decisions and 6 incorrect decisions), so a one-dimensional measurement cannot distinguish between subjects that scored 75% by making the same pattern of errors and subjects that scored 75% by making different patterns of errors.

To determine each chick's ceiling preference on the task, we measured the amount of time the chicks spent in proximity to their imprinted object during the rest periods. During the rest periods, the chicks could spend time with either the imprinted object or a white screen, so the rest periods provide a measure of the amount of time the chicks generally wanted to spend in proximity to their imprinted object (given that the chicks also spent time performing other behaviors such as eating and sleeping). We performed two complementary high-dimensional analyses using these coordinates.

First, we tested whether the chicks built similar representations as one another by measuring the Euclidean distance between their representations in the 24-dimensional space (Fig. 5A). To determine how close the representations were within the space, we computed the distance between the representations built by each pair of chicks. Each pair fell into one of four categories based on the chicks' imprinting stimulus. The pair of chicks could be imprinted to (1) the same view of the same object; (2) different views of the same object; (3) the same view of different objects; or (4) different views of different objects. In Experiment 1, the average representational distance between the chicks imprinted to the same view of the same object was not significantly different from the average distance between the chicks imprinted to the different views of the same object ($t(130) = 1.1$, $p = .28$, Cohen's $d = 0.19$). Conversely, the average representational distance between the chicks imprinted to the same view of different objects was larger than the average distance between chicks imprinted to the same view of the same object ($t(117.8) = 9.2$, $p < 10^{-14}$, Cohen's $d = 1.56$) and different views of the same object ($t(121.8) = 8.3$, $p < 10^{-12}$, Cohen's $d = 1.38$).

Similarly, in Experiment 2, the average representational distance between the chicks imprinted to the same view of the same object was not significantly different from the average distance between chicks imprinted to different views of the same object ($t(88) = 0.1$, $p = .93$, Cohen's $d = 0.02$). Conversely, the average representational distance between the chicks imprinted to the same view of different objects was larger than the average distance between chicks imprinted to the same view of the same object ($t(83.2) = 6.9$, $p < 10^{-9}$, Cohen's $d = 1.43$) and different views of the same object ($t(80.1) = 7.3$, $p < 10^{-9}$, Cohen's $d = 1.46$).

In both Experiments 1 and 2, the chicks that were imprinted to different objects built representations that were far away from one another in the recognition space (i.e., highly dissimilar representations), whereas the chicks that were imprinted to the different views of the same object built representations that were close together in the recognition space. Moreover, the chicks that were imprinted to different views of the same object built representations that were as close to one another in the recognition space as the chicks that were imprinted to the same view of the same object. Thus, the chicks built similar object representations despite being reared with different object views.

For our second analysis, we used principal component analysis (PCA) to reduce the dimensionality of the 24-dimensional recognition space. In a high-dimensional space, many dimensions can be correlated (e.g., performance on Viewpoint 1 could be correlated with performance on Viewpoint 2). PCA uses a linear transformation of the data to extract independent variables underlying performance. If the chicks imprinted to the different views of the same object built similar representations as one another, then a single principal component (object identity) should account for most of the variance in recognition performance. However, if the chicks built different representations from the different views of the objects, then two principal components (object identity and view) should be needed to account for the variance in recognition performance.

As shown in the scree plot in Fig. 5B, a single principal component accounted for most of the variance in performance in both Experiments 1 and 2. To determine which variables drove this principal component, we graphed each chick's first principal component score as a function of the chick's imprinted object and imprinted view (Fig. 5C). Visual

inspection of the graphs shows that the first principal component score was impacted by the chicks' imprinted object, but not by the imprinted view. Thus, the identity of the object—but not the viewpoint of the object—influenced the first principal component (which accounted for the majority of the variance across the 24 test viewpoints). In sum, these high-dimensional analyses demonstrate that newborn chicks develop common object representations from different views of an object.

5. General discussion

How does object recognition emerge in newborn brains during the earliest stages of visual learning? To address this question, we used a controlled-rearing method to systematically manipulate the amount of visual object input provided to newborn chicks. We reared chicks for five days with a virtual object shown from a single viewpoint, a small (11.25°) viewpoint range, or a large (360°) viewpoint range. We then measured the amount of view-invariance achieved by each chick by testing their object recognition performance across 24 viewpoints. Three main findings emerged.

First, in all three rearing conditions, the chicks successfully recognized their imprinted object across novel viewpoints. Thus, newborn chicks can develop view-invariant object recognition from a single view of an object—a case of one-shot learning in newborn brains. The chicks also developed similar levels of view-invariance when the object rotated through a small (11.25°) and large (360°) viewpoint range, despite the large viewpoint range providing 32 times more unique views of the object. This suggests that newborn visual systems are remarkably efficient at building view-invariant object representations.

Second, the chicks created similar object representations from different views of an object, indicating that newborn brains can converge on common object representations from different sets of sensory inputs. This feat was particularly challenging in our study because the within-object image difference was *greater* than the between-object image difference (from both pixel-level and V1-level perspectives, Fig. 4B). If the chicks had relied solely on image-level similarity to recognize objects (e.g., view-dependent image matching), then the chicks reared with the same views of different objects should have built more similar object representations than the chicks reared with different views of the same object. The finding that performance depended on object identity—rather than object view—indicates that newborn brains can successfully convert retinal inputs into object-centric representations.

Third, recognition performance improved when the chicks were reared with larger numbers of object views. While the chicks in the three rearing conditions showed similar generalization performance across *novel* views, the chicks also developed enhanced recognition for *familiar* views, leading the chicks reared with the large viewpoint range to outperform the chicks reared with the small viewpoint range. Consequently, both view-invariant and view-based learning—which have been well documented in studies of human adults (e.g., Biederman, 1987; Peissig & Tarr, 2007)—appear to be present at the onset of vision, when newborn animals build their first visual object representation.

These results compliment a growing body of work showing that humans and other animals are capable of one-shot learning. For example, both children and adults can learn a new concept from just one or a handful of examples (e.g., Lake, Salakhutdinov, & Tenenbaum, 2015; Landau, Smith, & Jones, 1988; Xu & Tenenbaum, 2007). The discovery that newborn chicks can perform one-shot learning suggests that this ability does not require extensive post-natal experience in order to develop. One-shot learning appears to scaffold object recognition during the earliest stages of learning.

5.1. The origins and nature of object recognition

Vision science has made substantial progress understanding how mature visual systems recognize objects. For instance, a large body of

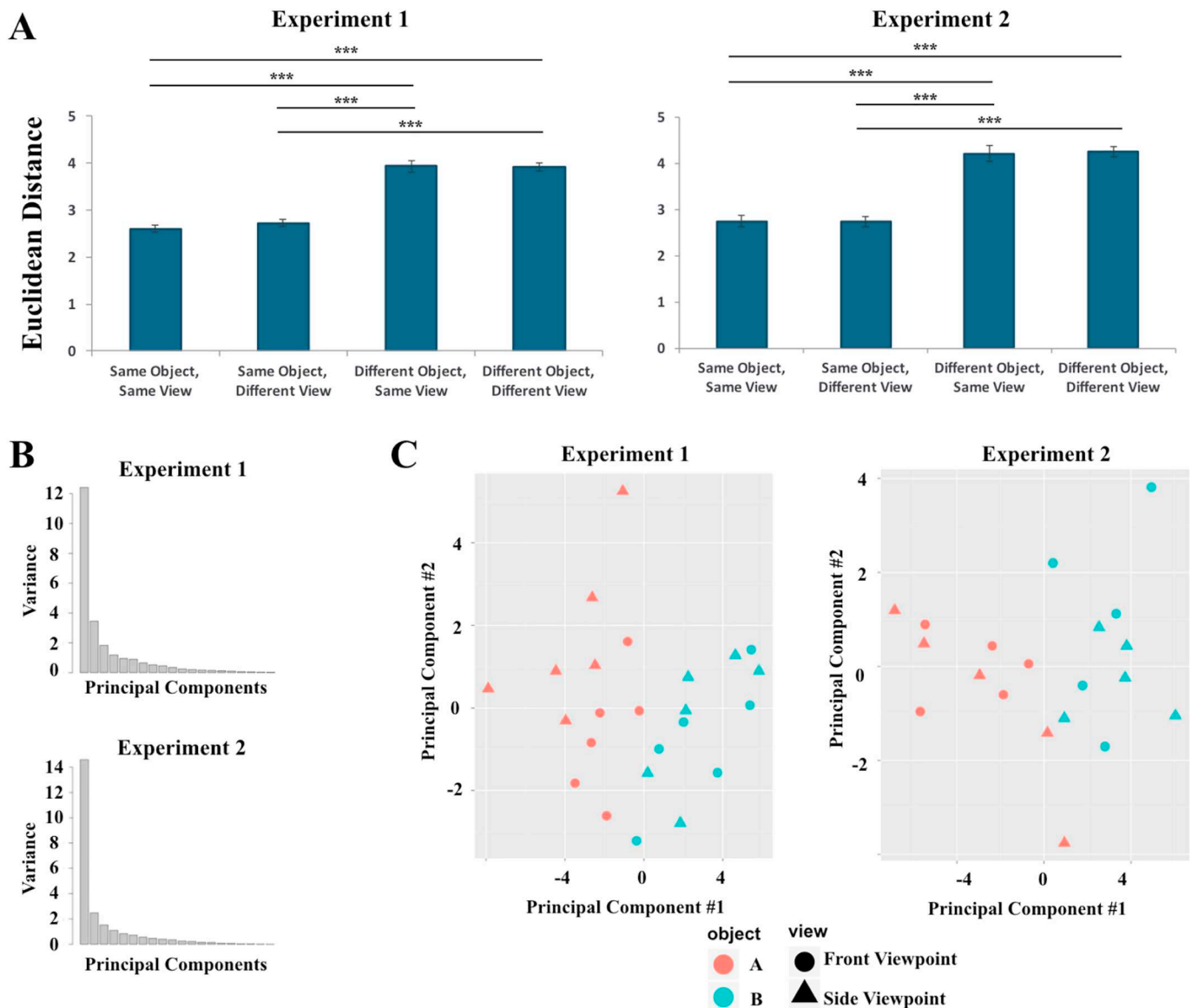


Fig. 5. High-dimensional analysis of recognition performance. (A) The average Euclidean distance between pairs of chicks in the 24-dimensional recognition space, as a function of the imprinted object and view. Each dimension of the recognition space reflects the chicks' performance for one of the 24 views. The chicks who were imprinted to the same object built representations that were closer together in the recognition space than the chicks who were imprinted to different objects. This effect was independent of the imprinted view. (B) Scree plots from the principal component analysis (PCA) in the 24-dimensional recognition space indicate one significant principal component. (C) The first principal component (explaining most of the variance) reflected the identity—rather than the viewpoint—of the object. Thus, newborn chicks successfully reformatted retinal inputs into object-centric representations that were selective for object identity and tolerant to viewpoint changes.

research provides evidence for both view-invariant and view-dependent recognition effects. Support for view-invariant recognition comes from studies showing that humans can recognize objects across large rotations in depth, size, and position (e.g., Biederman & Cooper, 1991, 1992; Biederman & Gerhardstein, 1993), while support for view-dependent recognition comes from studies showing enhanced recognition when objects are presented from familiar viewpoints (e.g., Tarr, Williams, Hayward, & Gauthier, 1998; Bulthoff, Edelman, & Tarr, 1995; Tarr, 1995). Our study extends this literature by showing that both view-invariant and view-dependent recognition effects are present when newborn animals build their first visual object representation.

The fact that chicks succeeded in a one-shot view-invariant recognition task—in the absence of all prior visual experience with objects—indicates that newborn brains can be highly generative. These

generative mechanisms might emerge from pre-natal experiences (e.g., spontaneous waves of retinal activity). A wave of retinal activation shares some characteristics with real-world objects: both are bounded, cohesive units that move on continuous spatiotemporal paths. Thus, early pre-natal experience might create initial priors for object recognition.

On the other hand, there is growing physiological and computational evidence that post-natal visual experience plays an important role in invariant object recognition (DiCarlo, Zoccolan, & Rust, 2012; Wallis & Bulthoff, 2001; Wiskott & Sejnowski, 2002). In fact, controlled-rearing studies provide direct evidence that experience with a natural visual world is essential for the development of invariant object recognition (Wood, 2016; Wood, Prasad, Goldman, & Wood, 2016; Wood & Wood, 2016, 2018). Specifically, newborn chicks require

experience with natural (slow and smooth) object motion in order to develop view-invariant recognition.⁴ When chicks are reared with objects that move too quickly, or do not move smoothly over time, the chicks fail to develop invariant recognition. Newborn brains appear to leverage natural (slow and smooth) visual inputs to learn object recognition. Human infants also rely on slow and smooth inputs during visual object learning (Smith & Slone, 2017), suggesting that this strategy might be a general characteristic of vertebrate visual development.

A different (but related) question concerns the nature of the view-invariant features the chicks used to succeed in this task. One possibility is that newborn chicks solve the invariance problem by building complex 3D representations of whole objects. Alternatively, chicks might learn to recognize objects by building representations of subfeatures that are smaller and less complex than the entire object. These subfeatures might respond to only a portion of the object, or be sensitive to key 2D, rather than 3D, features (e.g., Alemi-Neissi, Rosselli, & Zoccolan, 2013; Rosselli, Alemi, Ansuini, & Zoccolan, 2015). Indeed, leading computational models of invariant object recognition in primates explicitly rely on such subfeatures (Cadieu et al., 2014; Yamins et al., 2014; Yamins & DiCarlo, 2016).

Ultimately, understanding the origins and nature of object recognition will require a formal characterization of the visual processing machinery built during pre-natal development and how that machinery is shaped by experience. The present results should be useful for this enterprise because they provide detailed descriptions of how specific visual inputs relate to specific object recognition outputs in a newborn model system. These input-output patterns can serve as benchmarks for computational models. For example, in order to successfully account for the development of object recognition in newborn chicks, a computational model would need to produce three patterns. First, the model should be capable of building view-invariant object representations from sparse visual input (e.g., input of a single object seen from a single view) in an unsupervised manner. Second, the model should develop more accurate object representations when provided with natural (slow and smooth) visual object input. Third, the model should show both view-invariant and view-dependent learning effects.

5.2. Implications of these results for human visual development

To what extent do these results illuminate the development of human vision? From a neural circuit perspective, human and chicken sensory processing systems share strong similarities. Specifically, there is evidence that mammalian and avian brains process sensory input using a shared, evolutionarily-ancient cortical circuit that evolved over 300 million years ago in stem amniotes (Güntürkün & Bugnyar, 2016; Jarvis et al., 2005; Karten, 2013). The cortical circuits found in mammalian and avian brains are similar on the cellular, molecular, connectivity, and information-coding levels. If humans and chicks use homologous cortical circuits to process sensory input—as these findings suggest—then controlled-rearing experiments of newborn chicks can inform our understanding of human visual development.

Of course, there are many ways in which human and chicken brains differ from one another. Humans have much larger brains than chickens, which may influence both the amount of invariance that can

⁴ We hypothesize that newborn chicks' sensitivity to slow and smooth motion is related to spike-timing-dependent plasticity and the decay of neuronal firing over time. To illustrate, if objects move too quickly or non-smoothly, then greater numbers of retinal cells will be activated during the decay period, thereby transmitting a noisier signal to the next layer of neurons in the visual hierarchy. This hypothesis is consistent with our present results. In Experiment 1, the imprinted object moved slowly and smoothly over time. In Experiment 2, the view of the imprinted object appeared repeatedly at the same location, presumably causing the same population of retinal cells to fire repeatedly, without the noisy signal generated by fast and non-smooth object motion.

be obtained and the amount of time needed for development. Chickens—unlike humans—are also mobile from birth and immediately able to explore their environment. This active motor exploration likely plays an important role in the development of object perception (e.g., Smith & Slone, 2017; Soska, Adolph, & Johnson, 2010).

In conclusion, our study illuminates the origins of object recognition, revealing powerful one-shot learning that builds invariant representations from a single object view. These object representations can then be enriched through view-based learning, producing enhanced recognition for familiar views. These results reveal how object recognition emerges in newborn brains and provide high-precision benchmarks for building computational models of visual development.

Sample videos illustrating the design and stimuli are provided as supplementary materials. In addition, all data have been made publicly available via the Open Science Framework and can be accessed at <https://osf.io/f2sae/>. Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2020.104192>.

CRedit authorship contribution statement

Justin N. Wood: Conceptualization, Methodology, Data curation, Writing - original draft, Project administration, Funding acquisition. **Samantha M.W. Wood:** Conceptualization, Methodology, Data curation, Formal analysis, Writing - review & editing.

Acknowledgements

Funded by NSF CAREER Grant BCS-1351892 and a James S. McDonnell Foundation Understanding Human Cognition Scholar Award. This research was conducted at the University of Southern California.

References

- Alemi-Neissi, A., Rosselli, F. B., & Zoccolan, D. (2013). Multifetural shape processing in rats engaged in invariant visual object recognition. *Journal of Neuroscience*, *33*(14), 5939–5956.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147.
- Biederman, I., & Cooper, E. E. (1991). Evidence for complete translational and reflectional invariance in visual object priming. *Perception*, *20*(5), 585–593.
- Biederman, I., & Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 121.
- Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(6), 1162.
- Cadieu, C. F., Hong, H., Yamins, D. L. K., Pinto, N., Ardila, D., Solomon, E. A., et al. (2014). Deep neural networks rival the representation of primate IT cortex for core visual object recognition. *PLoS Computational Biology*, *10*(12), e1003963.
- DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. *Trends in Cognitive Sciences*, *11*(8), 333–341.
- DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, *73*(3), 415–434.
- Güntürkün, O., & Bugnyar, T. (2016). Cognition without cortex. *Trends in Cognitive Sciences*, *20*(4), 291–303.
- Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, *6*(2), 151–159.
- Karten, H. J. (1969). The organization of the avian telencephalon and some speculations on the phylogeny of the amniote telencephalon. *Annals of the New York Academy of Sciences*, *167*, 164–179.
- Karten, H. J. (2013). Neocortical evolution: Neuronal circuits arise independently of lamination. *Current Biology*, *23*(1), R12–R15.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 1097–1105.
- Lades, M., Vorbruggen, J. C., Buhmann, J., Lange, J., Vndermalsburg, C., Wurtz, R. P., et al. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers*, *42*, 300–311. <https://doi.org/10.1109/12.210173>.
- Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, *350*(6266), 1332–1338.
- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299–321.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

- Liu, G. B., & Pettigrew, J. D. (2003). Orientation mosaic in barn owl's visual Wulst revealed by optical imaging: Comparison with cat and monkey striate and extra-striate areas. *Brain Research*, 961(1), 153–158.
- Peissig, J. J., & Tarr, M. J. (2007). Visual object recognition: Do we know more now than we did 20 years ago? *Annual Review of Psychology*, 58(1), 75–96.
- Pettigrew, J. D., & Konishi, M. (1976). Neurons selective for orientation and binocular disparity in the visual Wulst of the barn owl (*Tyto alba*). *Science*, 193, 675–678.
- Rosselli, F. B., Alemi, A., Ansuini, A., & Zoccolan, D. (2015). Object similarity affects the perceptual strategy underlying invariant visual object recognition in rats. *Frontiers in Neural Circuits*, 9(197), 1–22 (10).
- Simonyan, K., & Zisserman, A. (2014). *Very deep convolutional networks for large-scale image recognition*. (arXiv:1409.1556).
- Smith, L. B., & Slone, L. K. (2017). A developmental approach to machine learning? *Frontiers in Psychology*, 8 5074–10.
- Soska, K. C., Adolph, K. E., & Johnson, S. P. (2010). Systems in development: Motor skill acquisition facilitates three-dimensional object completion. *Developmental Psychology*, 46(1), 129–138.
- Tarr, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review*, 2(1), 55–82.
- Tarr, M. J., Williams, P., Hayward, W. G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neuroscience*, 1(4), 275.
- Wallis, G., & Bühlhoff, H. H. (2001). Effects of temporal association on recognition memory. *Proceedings of the National Academy of Sciences*, 98(8), 4800–4804.
- Wiskott, L., & Sejnowski, T. J. (2002). Slow feature analysis: Unsupervised learning of invariances. *Neural Computation*, 14(4), 715–770.
- Wood, J. N. (2013). Newborn chickens generate invariant object representations at the onset of visual object experience. *Proceedings of the National Academy of Sciences*, 110(34), 14000–14005.
- Wood, J. N. (2014). Newly hatched chicks solve the visual binding problem. *Psychological Science*, 25(7), 1475–1481.
- Wood, J. N. (2015). Characterizing the information content of a newly hatched chick's first visual object representation. *Developmental Science*, 18(2), 194–205.
- Wood, J. N. (2016). A smoothness constraint on the development of object recognition. *Cognition*, 153, 140–145.
- Wood, J. N., Prasad, A., Goldman, J. G., & Wood, S. M. W. (2016). Enhanced learning of natural visual sequences in newborn chicks. *Animal Cognition*, 19(4), 835–845.
- Wood, J. N., & Wood, S. M. W. (2016). The development of newborn object recognition in fast and slow visual worlds. *Proceedings of the Royal Society B: Biological Sciences*, 283(1829).
- Wood, J. N., & Wood, S. M. W. (2017). Measuring the speed of newborn object recognition in controlled visual worlds. *Developmental Science*. <https://doi.org/10.1111/desc.12470>.
- Wood, J. N., & Wood, S. M. W. (2018). The development of invariant object recognition requires visual experience with temporally smooth objects. *Cognitive Science*, 39(17) 2885–16.
- Wood, S. M., & Wood, J. N. (2019). Using automation to combat the replication crisis: A case study from controlled-rearing studies of newborn chicks. *Infant Behavior and Development*, 57, 101329.
- Wood, S. M. W., & Wood, J. N. (2015). A chicken model for studying the emergence of invariant object recognition. *Frontiers in Neural Circuits*, 9(89), 7.
- Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, 114(2), 245–272.
- Xu, X., & Biederman, I. (2010). Loci of the release from fMRI adaptation for changes in facial expression, identity, and viewpoint. *Journal of Vision*, 10(14), 36.
- Yamins, D. L. K., & DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience*, 19(3), 356–365.
- Yamins, D. L. K., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences*, 111(23), 8619–8624.
- Zoccolan, D., Oertelt, N., DiCarlo, J. J., & Cox, D. D. (2009). A rodent model for the study of invariant visual object recognition. *Proceedings of the National Academy of Sciences*, 106(21), 8748–8753.