

Newborn chickens generate invariant object representations at the onset of visual object experience

Supplementary Information

Sections

1. Definition of object recognition and description of the virtual objects.....	2
2. Change in performance over time.....	6
3. Description of the hierarchical Bayesian analyses.....	7
4. Results of the Bayesian analyses for all experiments and viewpoint ranges.....	9
5. References.....	19

1. Definition of object recognition and description of the virtual objects

Following previous research (1), object recognition was defined as the ability to discriminate an object from other objects, and to do so over a range of identity-preserving transformations of the retinal image of that object (e.g., image transformations resulting from changes in viewpoint). This definition is useful for two reasons. First, it focuses attention on the computational hallmark of object recognition—the ability to identify objects over a large range of viewing conditions. Although each encounter with an object is almost entirely unique (in terms of the image projected onto the retina), the visual system is able to link these unique patterns of retinal activity to the same object stored in memory. This so-called “invariance problem” is the computational crux of object recognition (2), and the primary obstacle in the development of artificial object recognition systems (3). Second, by focusing on the invariance problem, this definition provides an objective measure of having “solved” object recognition. An organism can be said to have solved object recognition if it can recognize objects across large and complex changes in their pixel-level appearances.

The two virtual objects (Fig. S1) were modeled after those used in a previous study that tested for invariant object recognition in adult rats (4). These objects are ideal for studying invariant recognition because changing the viewpoint (i.e., azimuth and elevation rotation) of a given object can produce a greater within-object image difference than changing the identity of the object while maintaining its viewpoint. Thus, distinguishing between these objects from novel viewpoints requires an invariant representation that can generalize across large, novel, and complex changes in the object’s pixel-level appearance.

To confirm that subjects were able to recognize their imprinted object across large and complex changes in its pixel-level appearance, I measured the amount of image variation produced by the virtual objects. A retinal image consists of a collection of signals from photoreceptor cells. Each photoreceptor cell registers the brightness value from a particular region in the image. Thus, it is possible to regard a retinal image as a matrix of numbers, each specifying light intensity at an image pixel (5). To compare the pixel-wise similarity of the virtual objects, I converted each object animation into a sequence of images (at 24 frames per second) and measured the brightness level of each pixel in each image. For each corresponding set of frames (e.g., the 1st frame of the input animation and the 1st frame of the test animation, the 2nd frame of the input animation and the 2nd frame of the test animation, etc.), I compared the brightness level of each corresponding pixel. I then added together the absolute differences in brightness values obtained for all corresponding pixels, across all corresponding frames, to obtain a single measure of the overall pixel difference between the test animation and the input animation. This analysis confirmed that changing the viewpoint of the imprinted object could produce a larger pixel-wise image change than manipulating the identity of the object while maintaining its viewpoint. As shown in Figures S2 and S3, the within-object image difference (i.e., the pixel-level difference between the test animation of the imprinted object and the input animation of the imprinted object) was greater than the between-object image difference (i.e., the pixel-level difference between the test animation of the unfamiliar object and the input animation of the imprinted object) on 15 of the 22 novel viewpoint tests in Experiment 1, and 22 of the 22 novel viewpoint tests in Experiment 2. These results supplement the analyses of Zoccolan and colleagues (4), who also measured the amount of image variation produced by changing the viewpoint and

identity of these two objects. Importantly, they showed that when these objects are presented from the same azimuth and elevation positions, the between-object image difference is smaller than the within-object image distance obtained when the same object is presented from different azimuth and elevation positions. These researchers measured image variation in several ways, including (a) computing the pixel-wise Euclidean distance between the different images of the objects, and (b) computing image difference by using the responses of a population of simulated V1-like simple cells (simulated using a bank of Gabor filters). These two measurements yielded similar within- and between-object image differences (4).

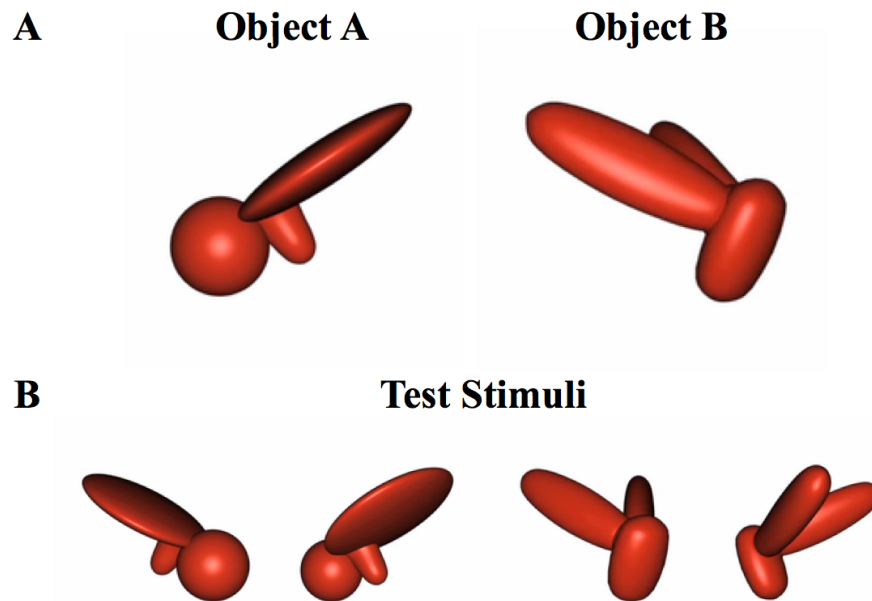


Figure S1: (A) The two virtual objects used in the study. (B) When presented with four novel test stimuli of these objects, a human observer can easily recognize the test stimuli as being “Object A, Object A, Object B, Object B” (from left to right). In contrast, an ideal observer having access to all available image information, but lacking knowledge of the geometry of identity-preserving transformations (i.e., a classification based on a pixel-by-pixel comparison), would incorrectly judge the test stimuli to be “Object B, Object A, Object B, Object A” (example adapted from ref. 9).

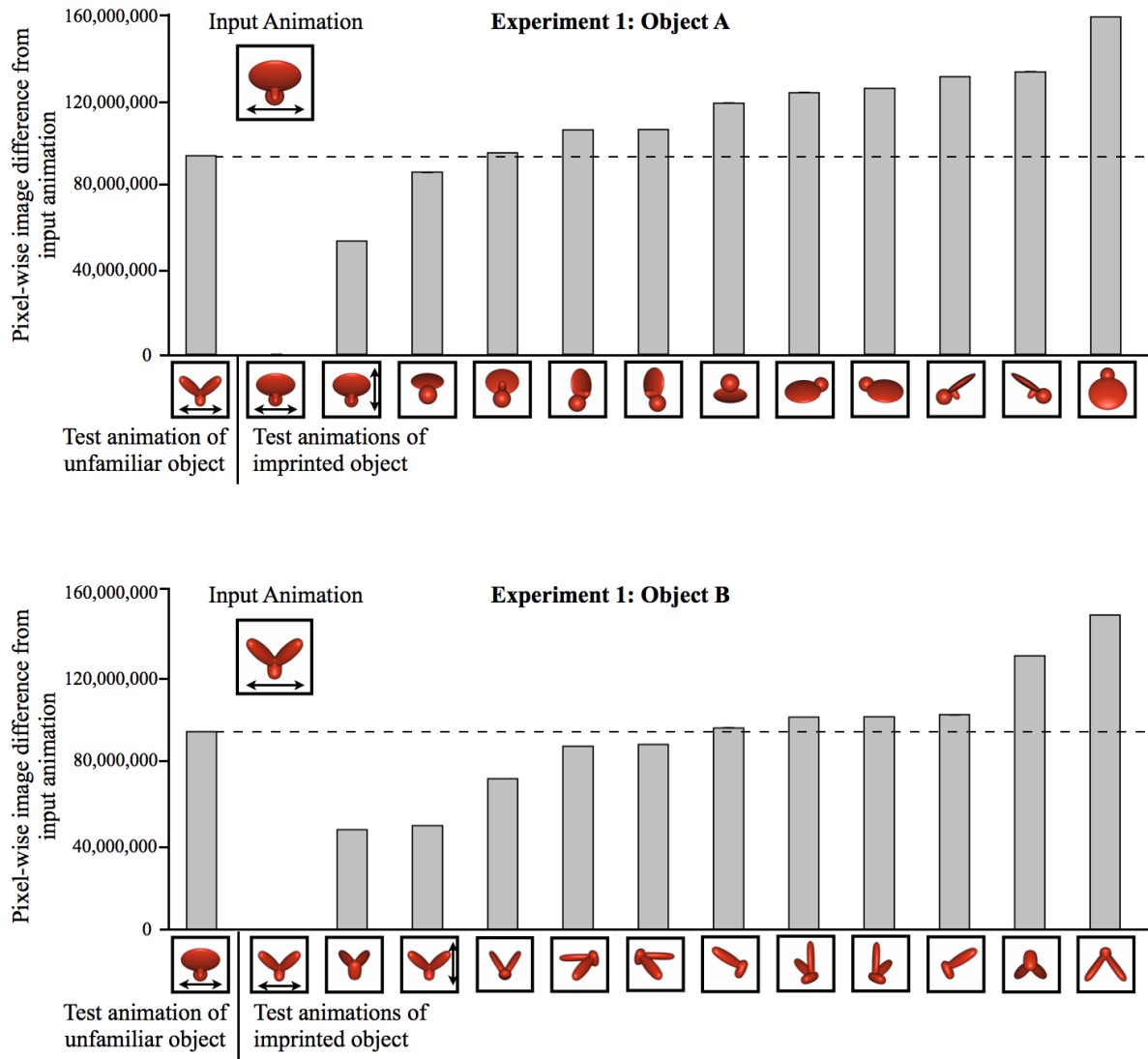


Figure S2: Results of the pixel-by-pixel comparisons of the virtual objects presented in the test phase and the virtual object presented in the input phase for Experiment 1. The dashed line shows the image difference between the unfamiliar object and imprinting stimulus (i.e., the input animation). Critically, for most of the test trials, the unfamiliar object was more similar to the imprinting stimulus than the imprinted object was to the imprinting stimulus.

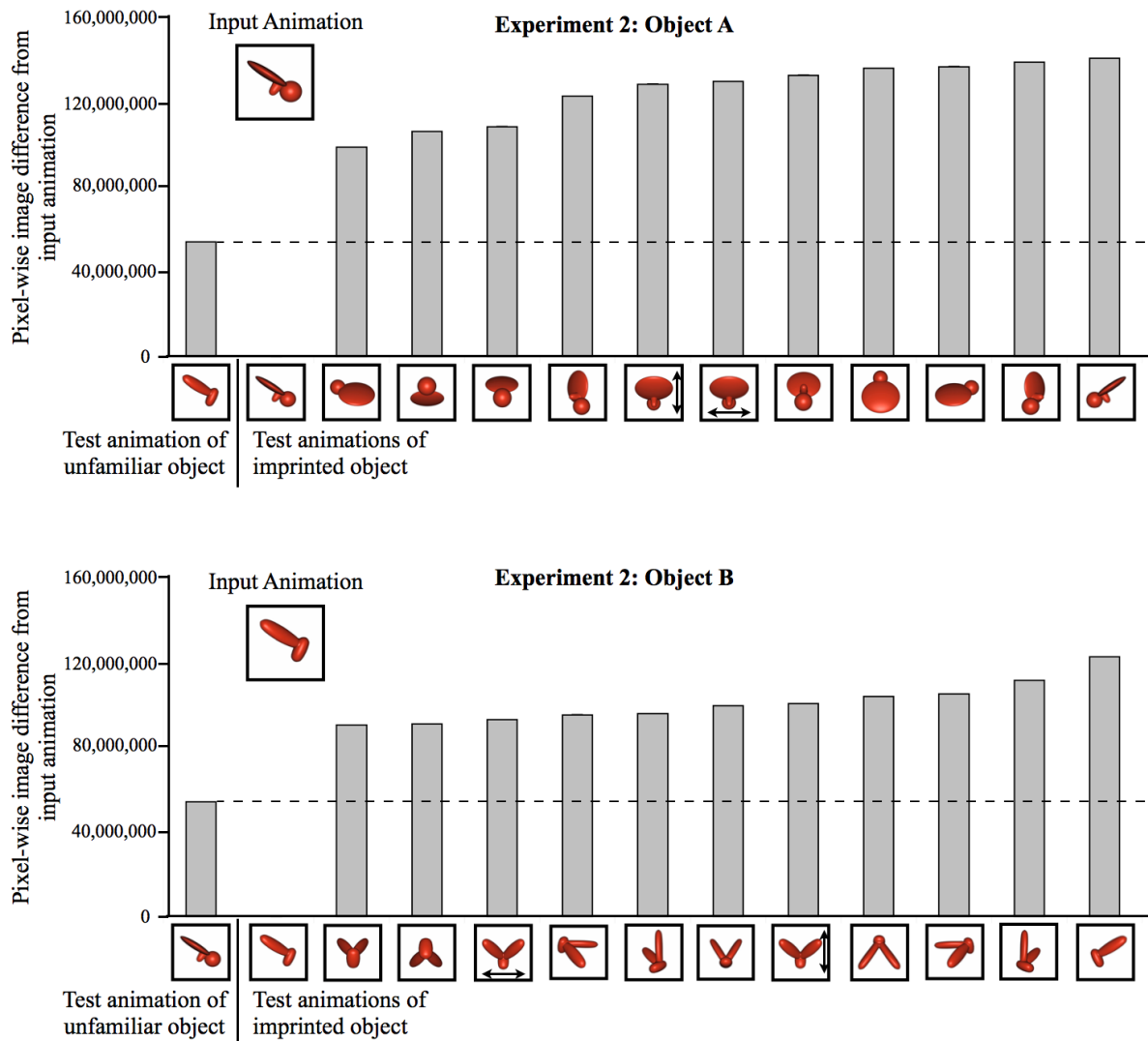


Figure S3: Results of the pixel-by-pixel comparisons of the virtual objects presented in the test phase and the virtual object presented in the input phase for Experiment 2. The dashed line shows the image difference between the unfamiliar object and imprinting stimulus (i.e., the input animation). Critically, for all of the novel viewpoint test trials, the unfamiliar object was more similar to the imprinting stimulus than the imprinted object was to the imprinting stimulus.

2. Change in performance over time

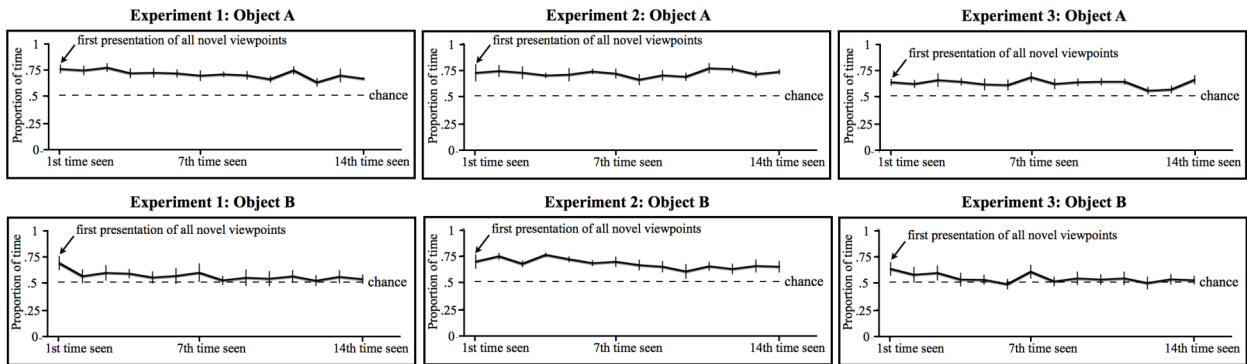


Figure S4: Change over time results. The graphs illustrate group mean performance over the full set of viewpoint ranges shown during the test phase, computed for the first, second, third, etc., presentation of the viewpoint ranges. The y-axis indicates the proportion of time subjects spent with the imprinted object versus the unfamiliar object. Chance performance was 50%. Error bars denote standard error.

3. Description of the hierarchical Bayesian analyses

To analyze the data, I first computed the number of test trials in which subjects preferred their imprinted object over the unfamiliar object. The subject was rated to have preferred their imprinted object on a trial if their object preference score was greater than 50%. The object preference score was calculated with the formula:

$$\text{Object Preference Score} = \frac{\text{Time by Imprinted Object}}{\text{Time by Imprinted Object} + \text{Time by Unfamiliar Object}}$$

Test trials were thus scored as “correct” when subjects spent a greater proportion of time with their imprinted object and “incorrect” when they spent a greater proportion of time with the unfamiliar object. These responses were then analyzed using hierarchical Bayesian methods (6) that are able to account for the hierarchical dependencies in the data (shown in Fig. S5).

Bayesian methods offer many advantages over traditional null hypothesis significance testing (7). For instance, with a Bayesian analysis, it is possible to calculate the actual probability that performance was above chance levels (an intuitive statistic to interpret), rather than a p -value that calculates the probability of getting data as extreme as the data actually obtained assuming that the null hypothesis is actually true (a less intuitive statistic to interpret). For the current

Hierarchical Dependencies in the Data

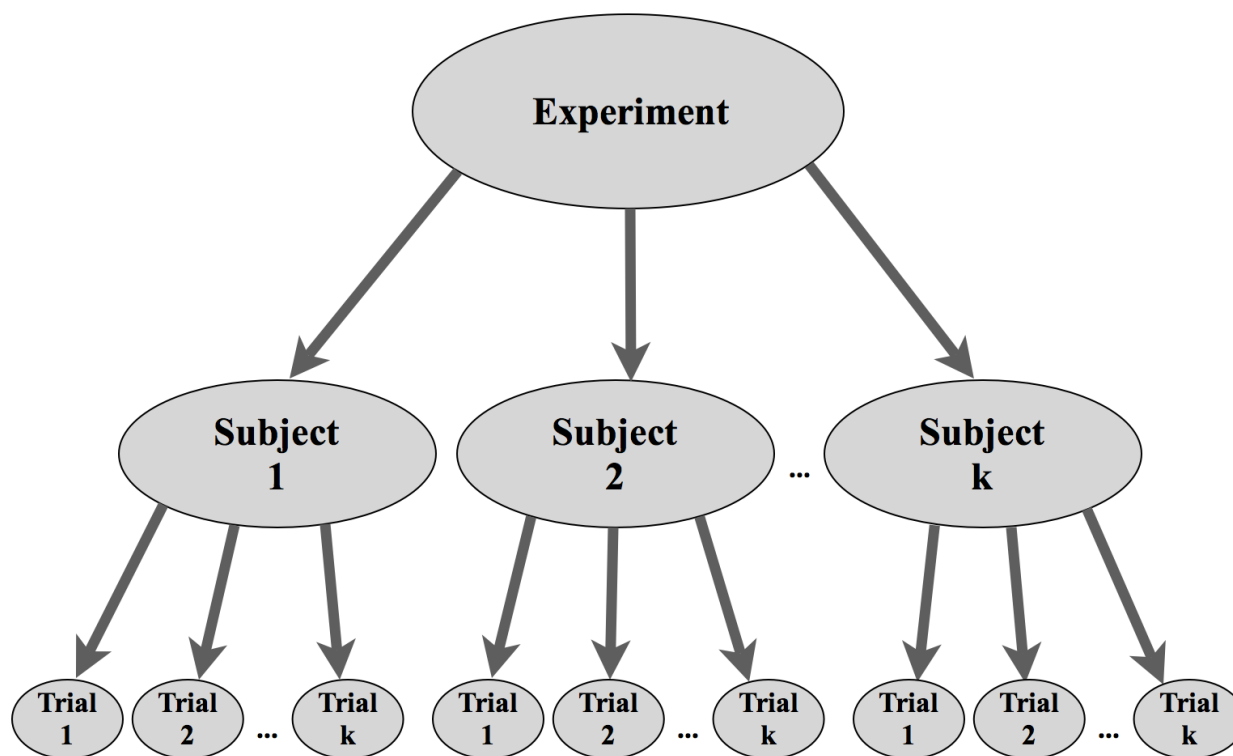


Figure S5: A Bayes net diagram depicting the hierarchical dependency of the data.

study, Bayesian analyses provided detailed probabilistic estimates of success for each newborn subject as well as the underlying probability of success for each condition and experiment (for results, see section 4).

This Bayesian analysis first requires specifying a prior distribution that represents the current uncertainty in performance. I used a prior consisting of one correct trial and one incorrect trial. This is a highly conservative prior because it consists of just two trials at chance level (50%). The prior distribution also includes a parameter, K , that represents the consistency across subjects. To be conservative, I used a uniform prior (8) that ranged from 0.000001 (i.e., very little consistency across subjects) to the maximum reasonable kappa. The maximum reasonable kappa was estimated from subjects' performance during the rest periods in the test phase. During the rest periods, the input animation from the input phase was projected onto one display wall and a white screen was projected onto the other display wall. These rest periods were expected to produce the greatest consistency across subjects because they presented the easiest choice: subjects chose whether to spend time with the imprinting stimulus versus a white screen. Critically, all of these prior distribution parameters were both broad and vague, thereby expressing great prior uncertainty in the values of the parameters. This type of prior has minimal influence on the estimates of success and is quickly overwhelmed by even a modest amount of data with Bayesian parameter estimation (6).

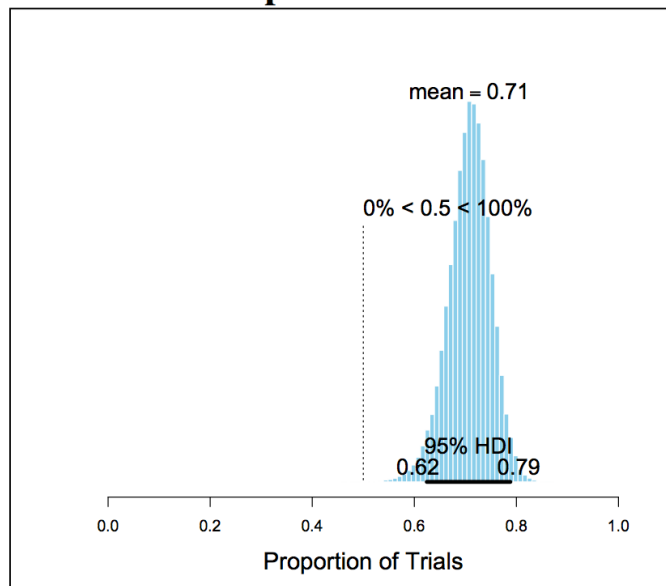
All analyses were performed using R version 2.15.0 (<http://www.r-project.org/>), JAGS (<http://mcmc-jags.sourceforge.net/>), and adaptation of code from Dr. John Kruschke (6). The hierarchical Bayesian model used Markov Chain Monte Carlo (MCMC) sampling to approximate the posterior distribution of the parameters for each individual subject and the hyperparameter for each condition. The analysis used a burn-in of 10,000 steps, with a total of 100,000 steps after burn-in.

4. Results of the Bayesian analyses for all experiments and viewpoint ranges

This section provides a graphical depiction of performance for each experiment and viewpoint test. Figures S6-S8 provide the probability density graphs for the group and the probability density graphs for each subject. Figures S9-S14 provide the probability density graphs for the group for all of the viewpoint tests in Experiments 1-3. The graphs show an estimate of the 95% highest density interval (HDI). The 95% HDI is an interval that spans 95% of the distribution, such that every point inside the interval has higher believability than any point outside of the interval (6). Chance performance (50%) is indicated by the vertical dashed line on each graph.

Experiment 1

Group Performance



Subject Performance

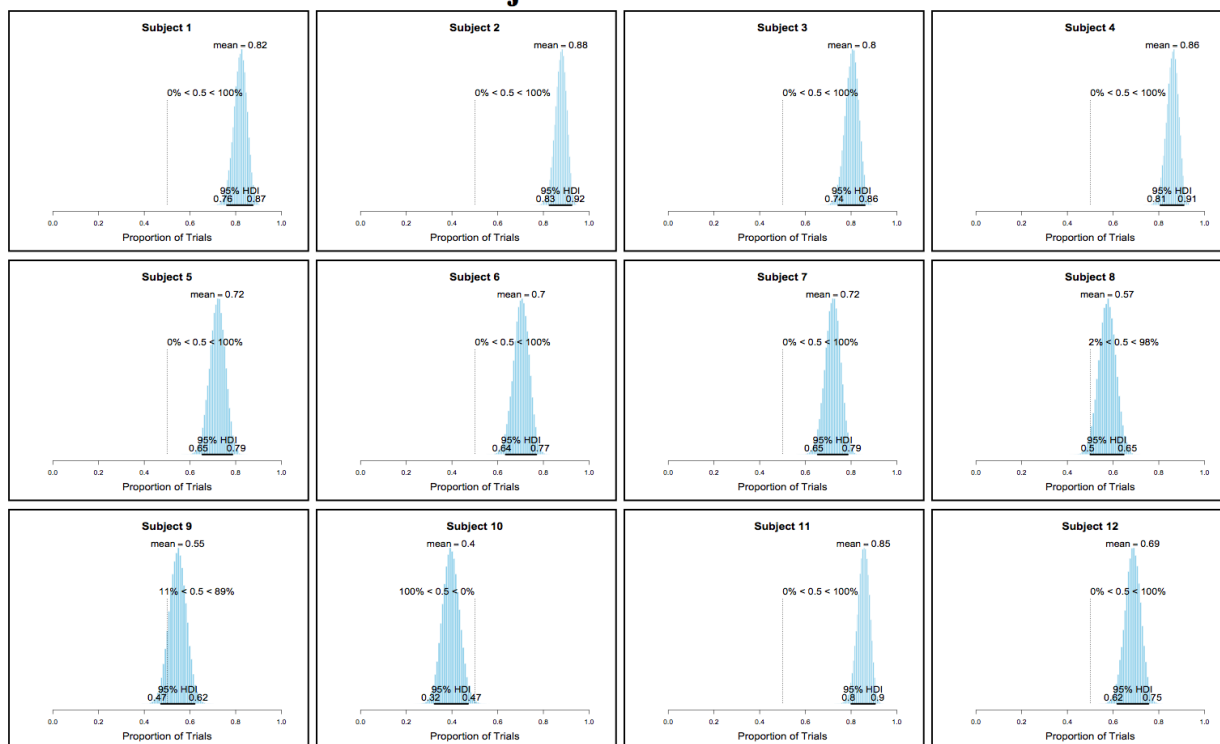
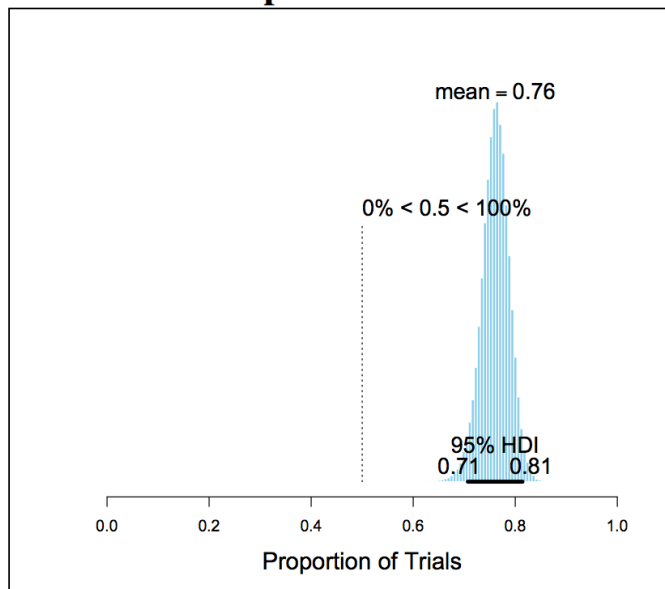


Figure S6: The probability density graphs for the group and each subject in Experiment 1.

Experiment 2

Group Performance



Subject Performance

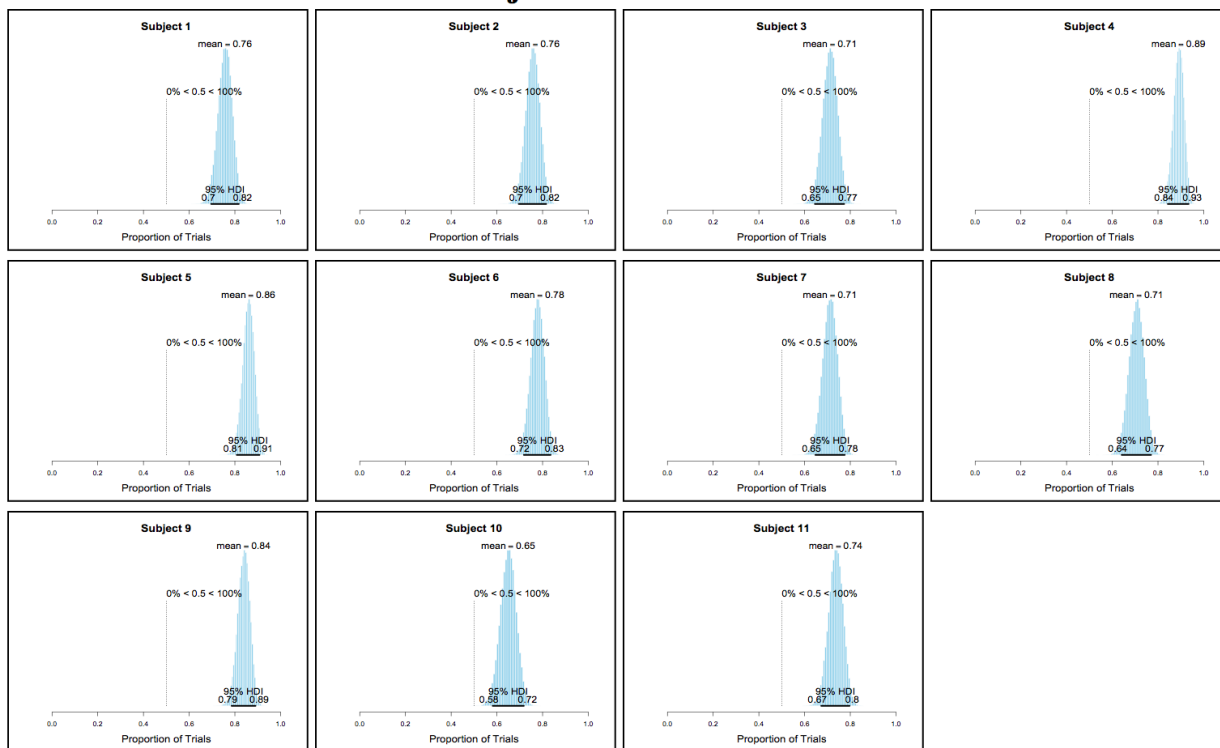
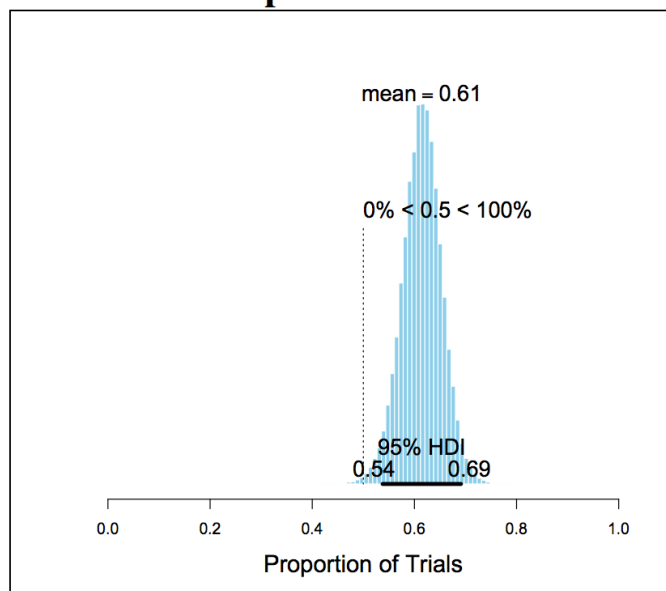


Figure S7: The probability density graphs for the group and each subject in Experiment 2.

Experiment 3

Group Performance



Subject Performance

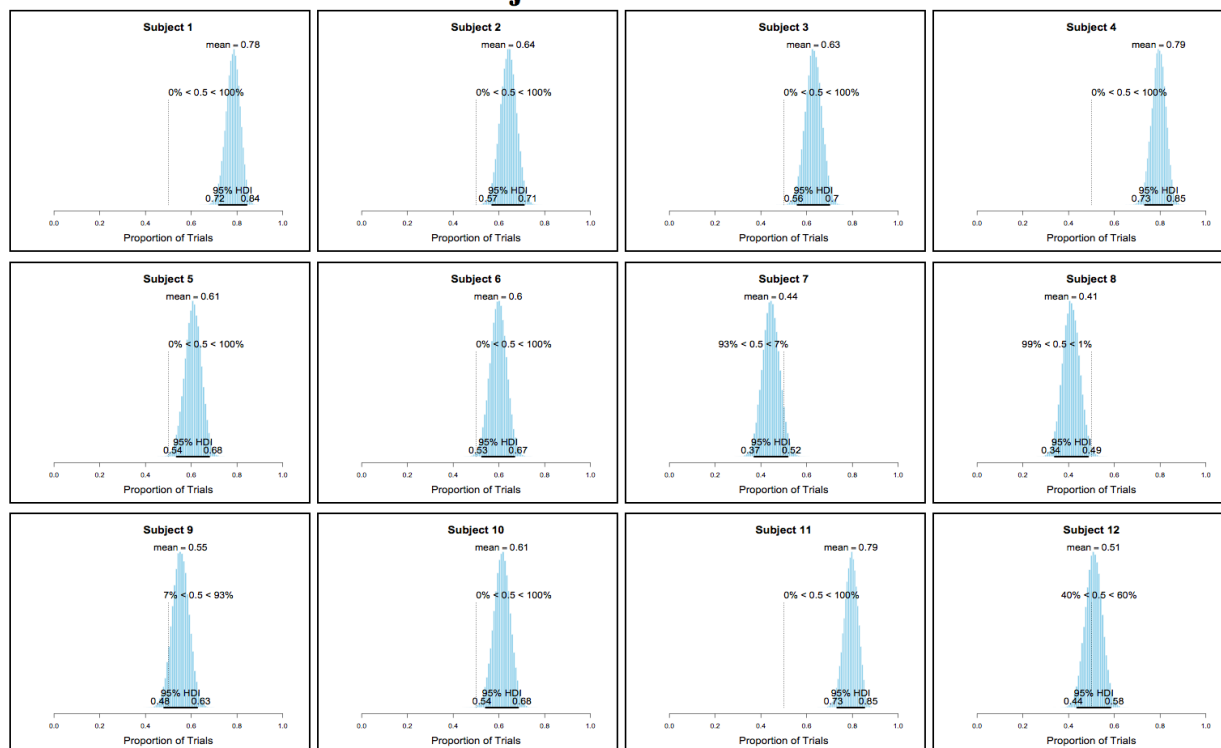


Figure S8: The probability density graphs for the group and each subject in Experiment 3.



Experiment 1: Imprinted to frontal viewpoint range of Object A

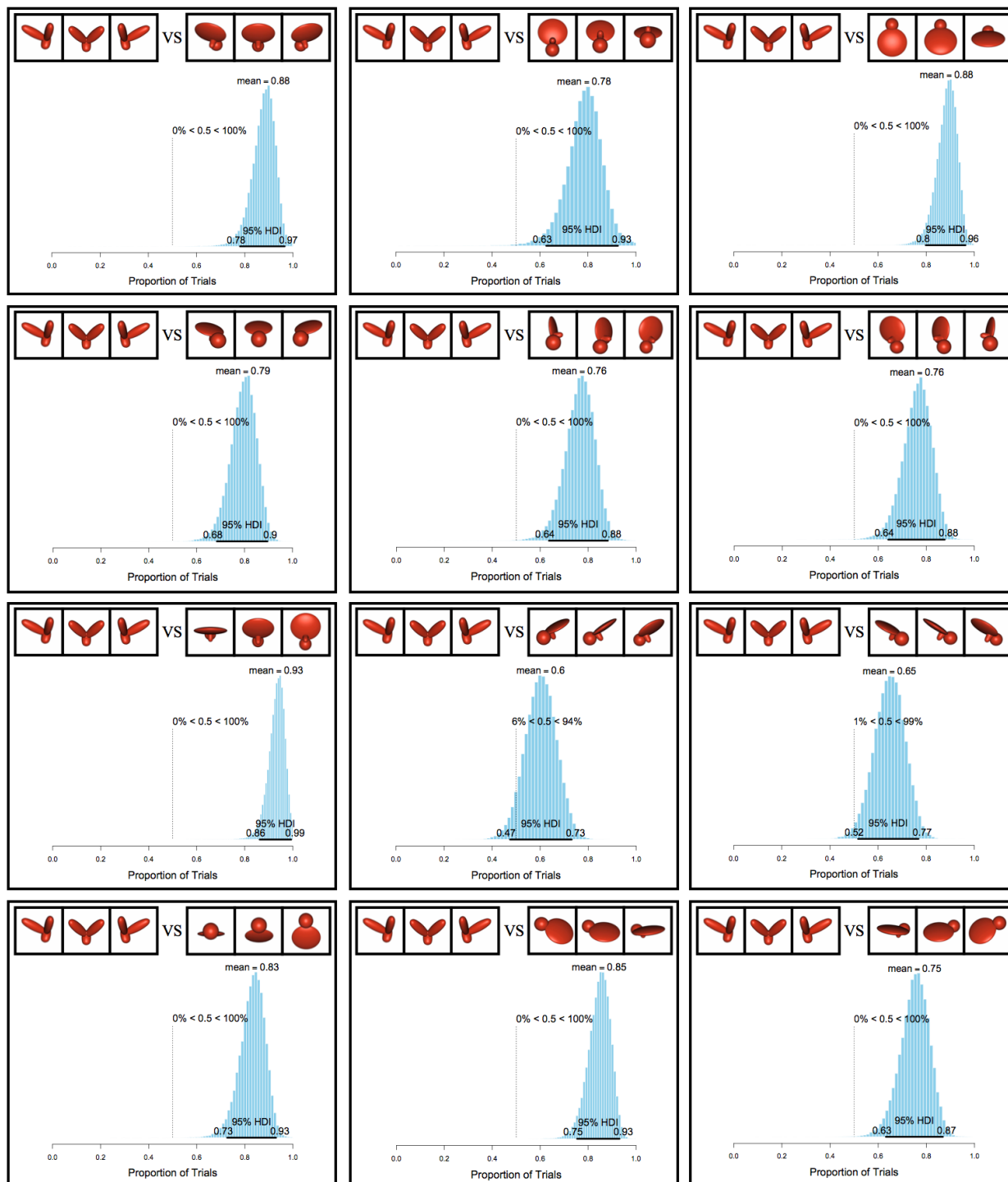


Figure S9: The probability density graphs for the group for each viewpoint test in Experiment 1. These subjects were imprinted to Object A.



Experiment 1: Imprinted to frontal viewpoint range of Object B

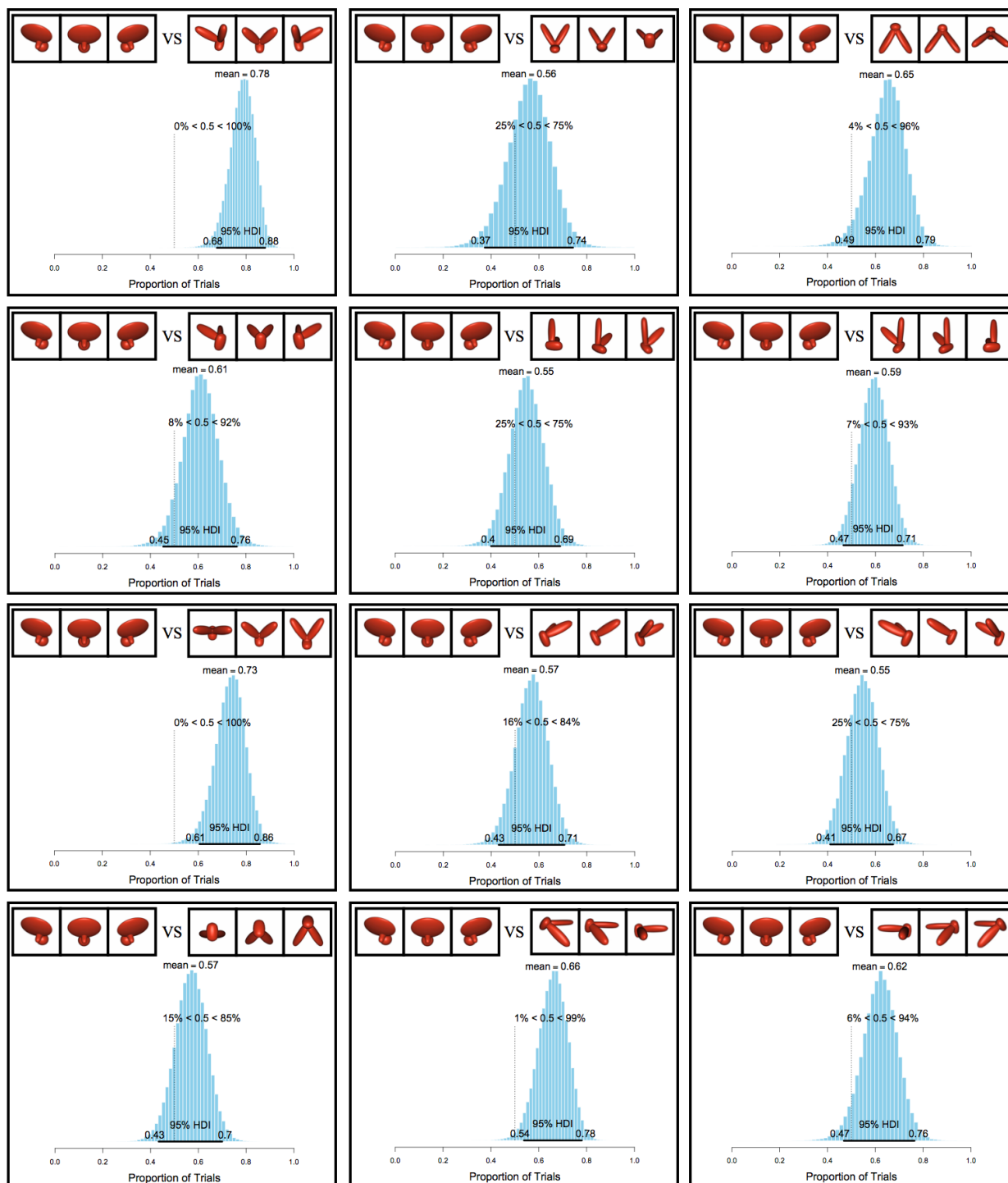


Figure S10: The probability density graphs for the group for each viewpoint test in Experiment 1. These subjects were imprinted to Object B.



Experiment 2: Imprinted to side viewpoint range of Object A

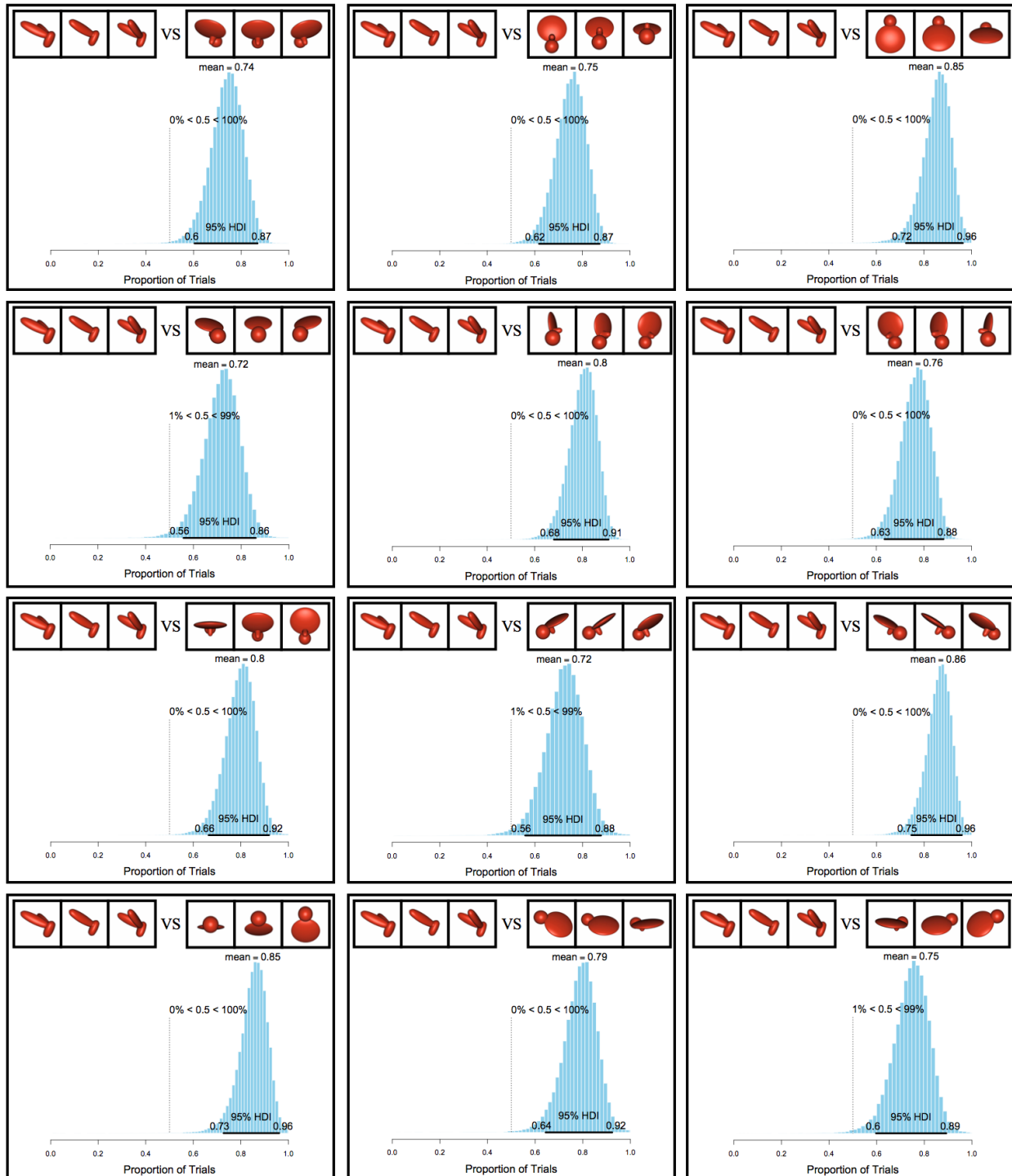


Figure S11: The probability density graphs for the group for each viewpoint test in Experiment 2. These subjects were imprinted to Object A.



Experiment 2: Imprinted to side viewpoint range of Object B

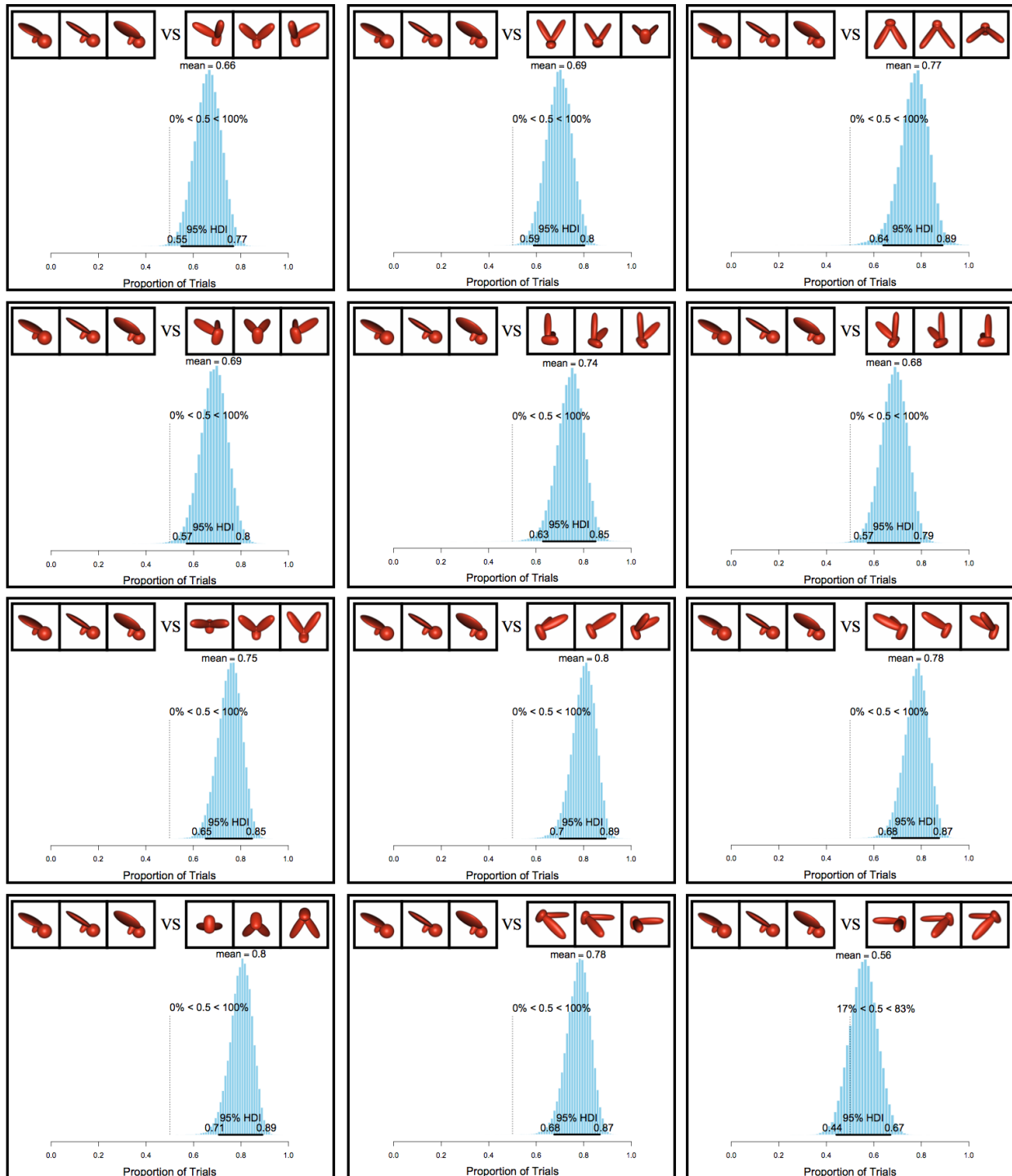


Figure S12: The probability density graphs for the group for each viewpoint test in Experiment 2. These subjects were imprinted to Object B.


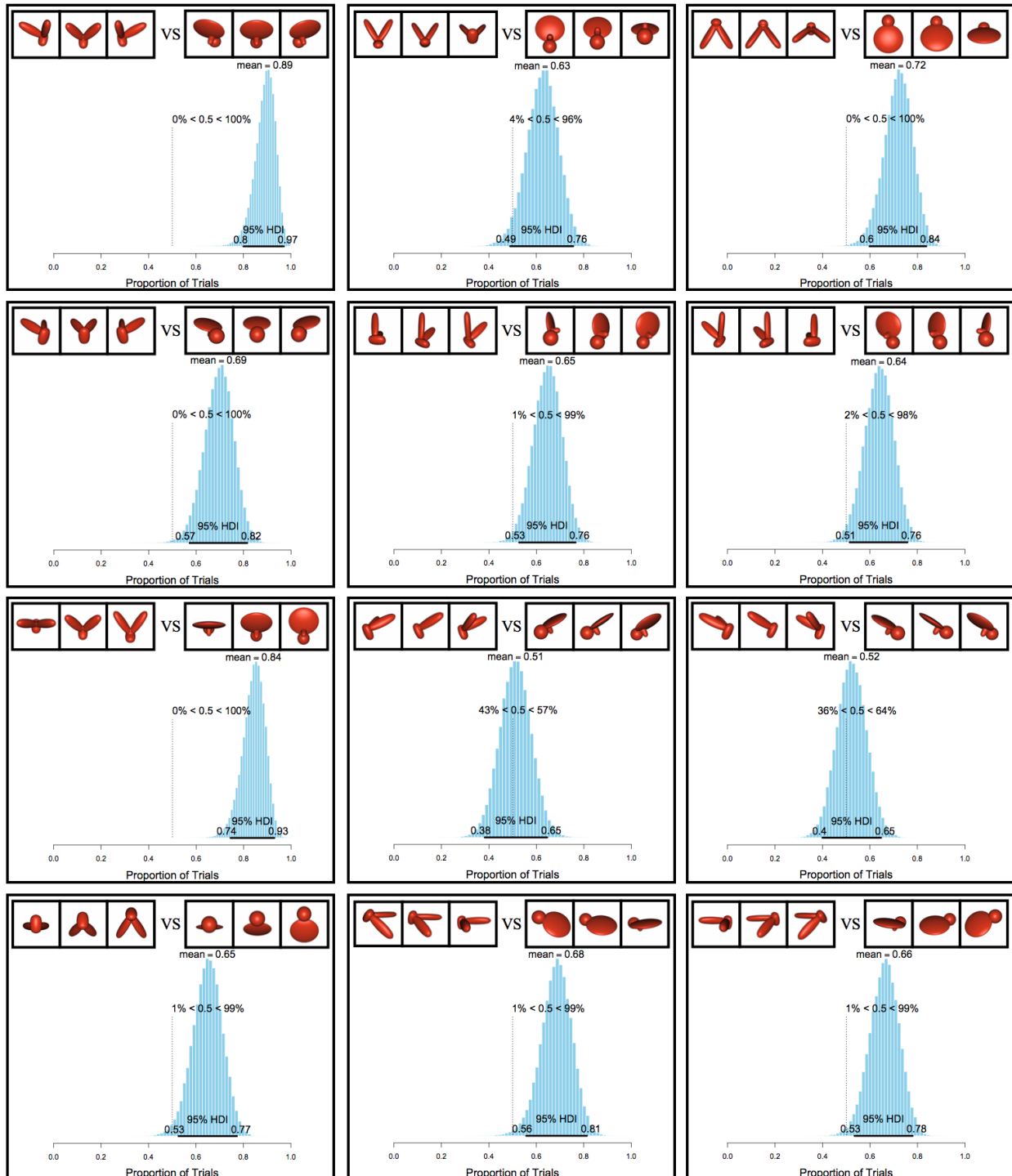

Experiment 3: Imprinted to frontal viewpoint range of Object A


Figure S13: The probability density graphs for the group for each viewpoint test in Experiment 3. These subjects were imprinted to Object A.



Experiment 3: Imprinted to frontal viewpoint range of Object B

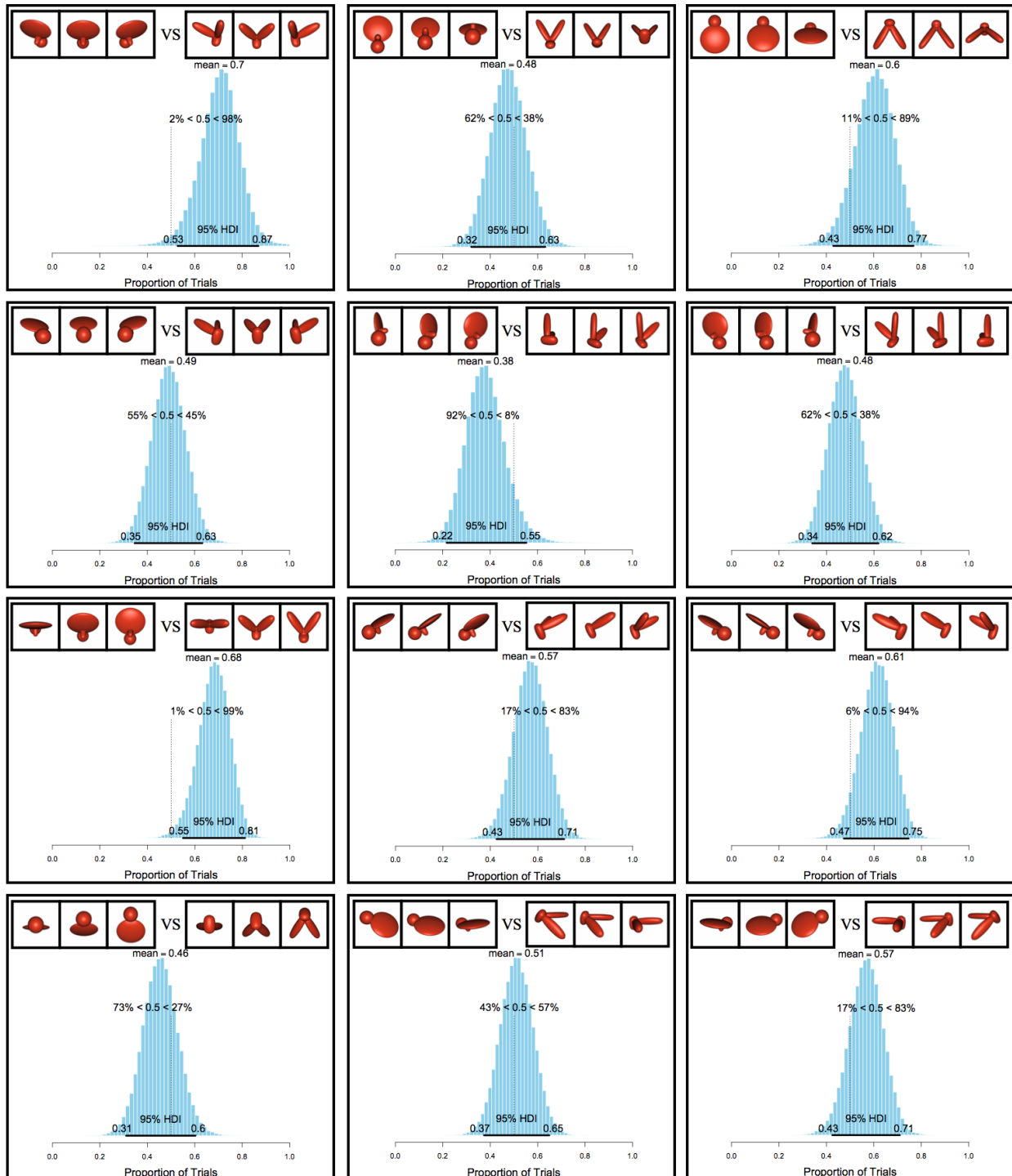


Figure S14: The probability density graphs for the group for each viewpoint test in Experiment 3. These subjects were imprinted to Object B.

5. References

1. DiCarlo JJ & Cox DD (2007) Untangling invariant object recognition. *Trends in Cognitive Sciences* 11(8):333-341.
2. DiCarlo JJ, Zoccolan D, & Rust NC (2012) How does the brain solve visual object recognition? *Neuron* 73(3):415-434.
3. Pinto N, Cox DD, & DiCarlo JJ (2008) Why is Real-World Visual Object Recognition Hard? *PLoS Comput Biol* 4(1):e27.
4. Zoccolan D, Oertelt N, DiCarlo JJ, & Cox DD (2009) A rodent model for the study of invariant visual object recognition. *Proc Natl Acad Sci USA* 106(21):8748-8753.
5. Seung HS & Lee DD (2000) The manifold ways of perception. *Science* 290(5500):2268-2269.
6. Kruschke JK (2011) *Doing bayesian data analysis: a tutorial with R and BUGS* (Academic Press, Burlington, MA).
7. Kruschke JK (2010) What to believe: Bayesian methods for data analysis. *Trends in Cognitive Sciences* 14(7):293-300.
8. Gelman A (2006) Prior distributions for variance parameters in hierarchical models. *Bayesian Anal* 1(3):515-533.
9. Goris RLT & Op de Beeck HP (2010). Invariance in visual object recognition requires training: a computational argument. *Frontiers in Neuroscience*. doi: 10.3389/neuro.01.012.2010