Set of images

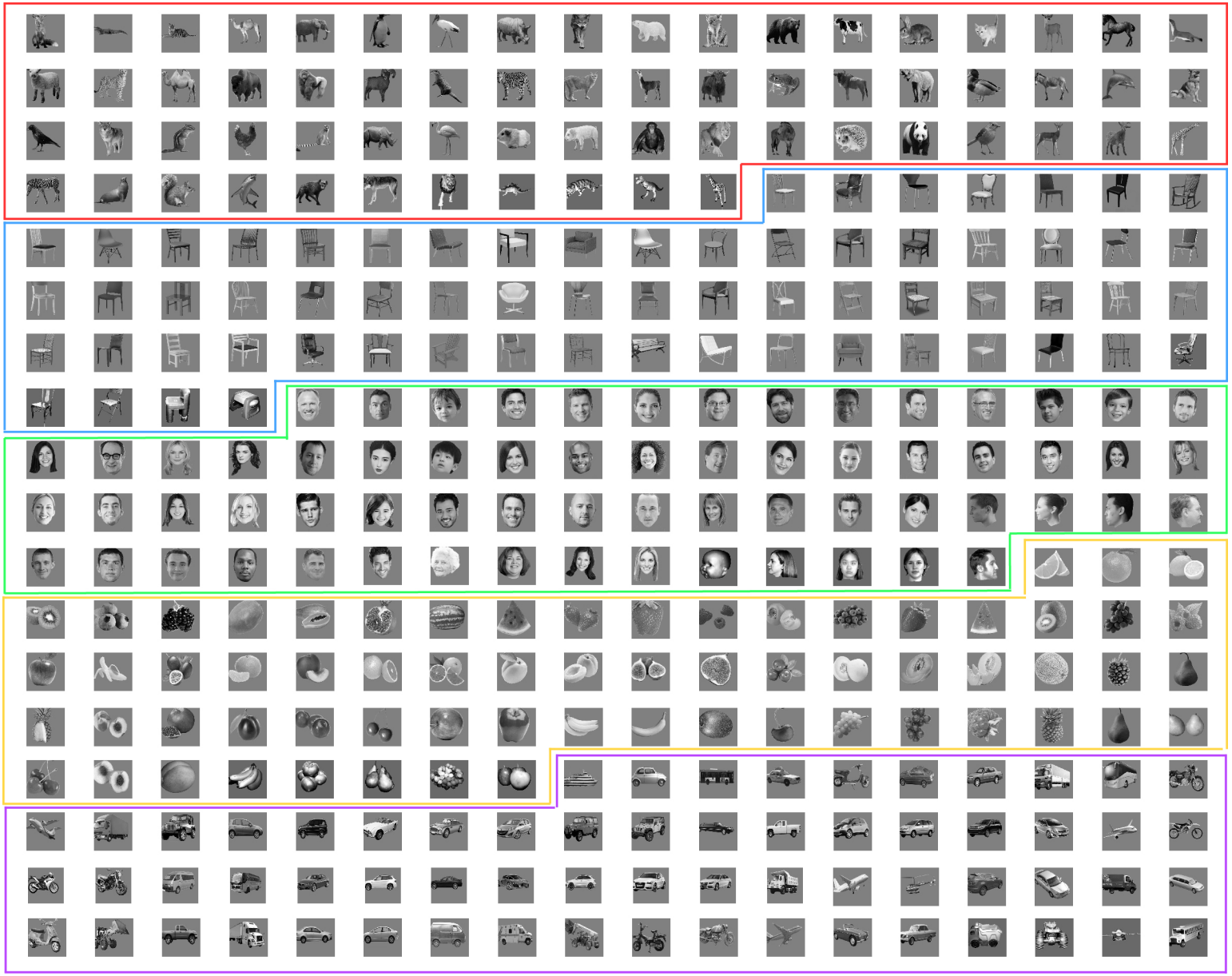

Figure W1: Set of images used in the experiments
The bounding boxes separate the 5 different categories

Pre-trained feed-forward models performed below humans

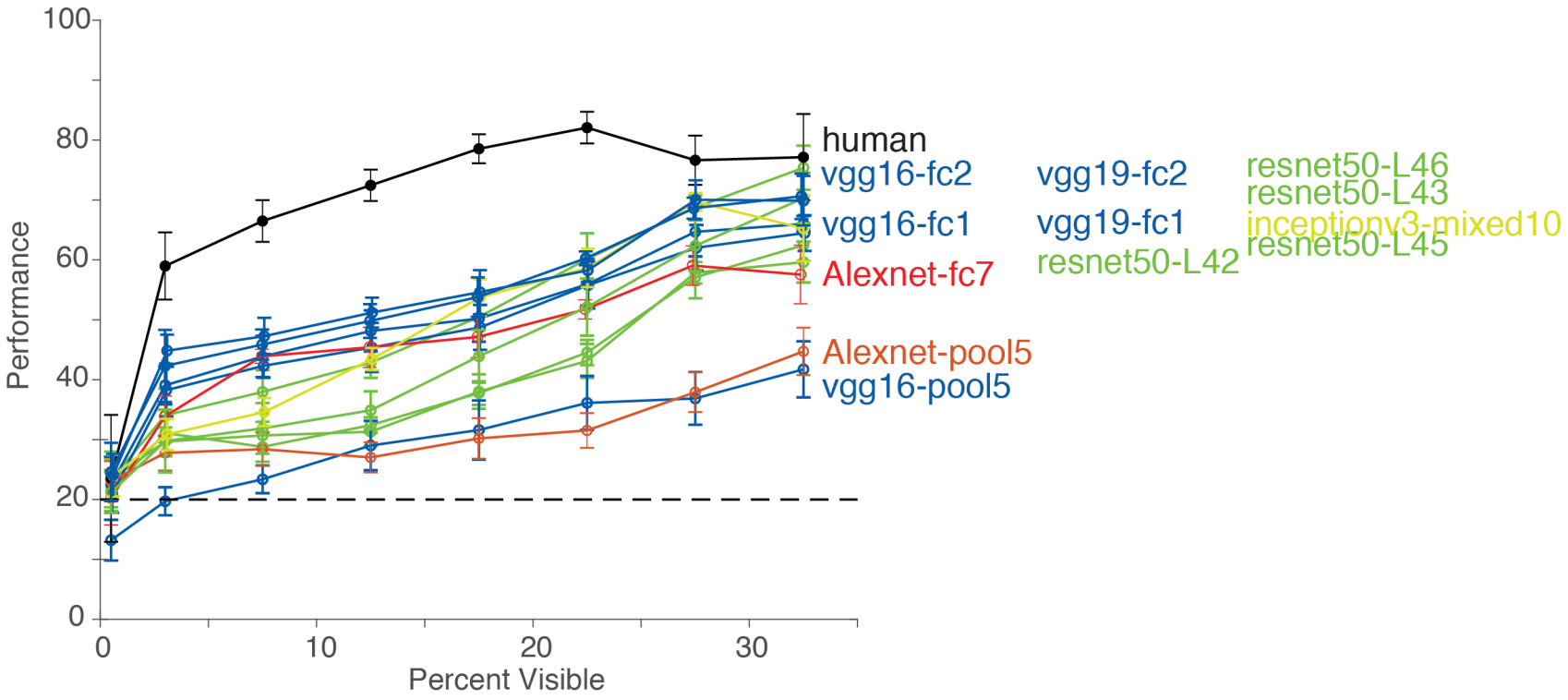

Figure W2: None of the feed-forward models tested achieved human-level performance under low-visibility conditions

Expanding on **Figure S4**, we evaluated multiple different processing stages in some of the top performing feed-forward models for visual recognition (Alexnet, VGG16, VGG19, Resnest, Inception). The format and conventions for this figure follow those in **Figure S4**.

Adding recurrent connections improved recognition performance of occluded objects

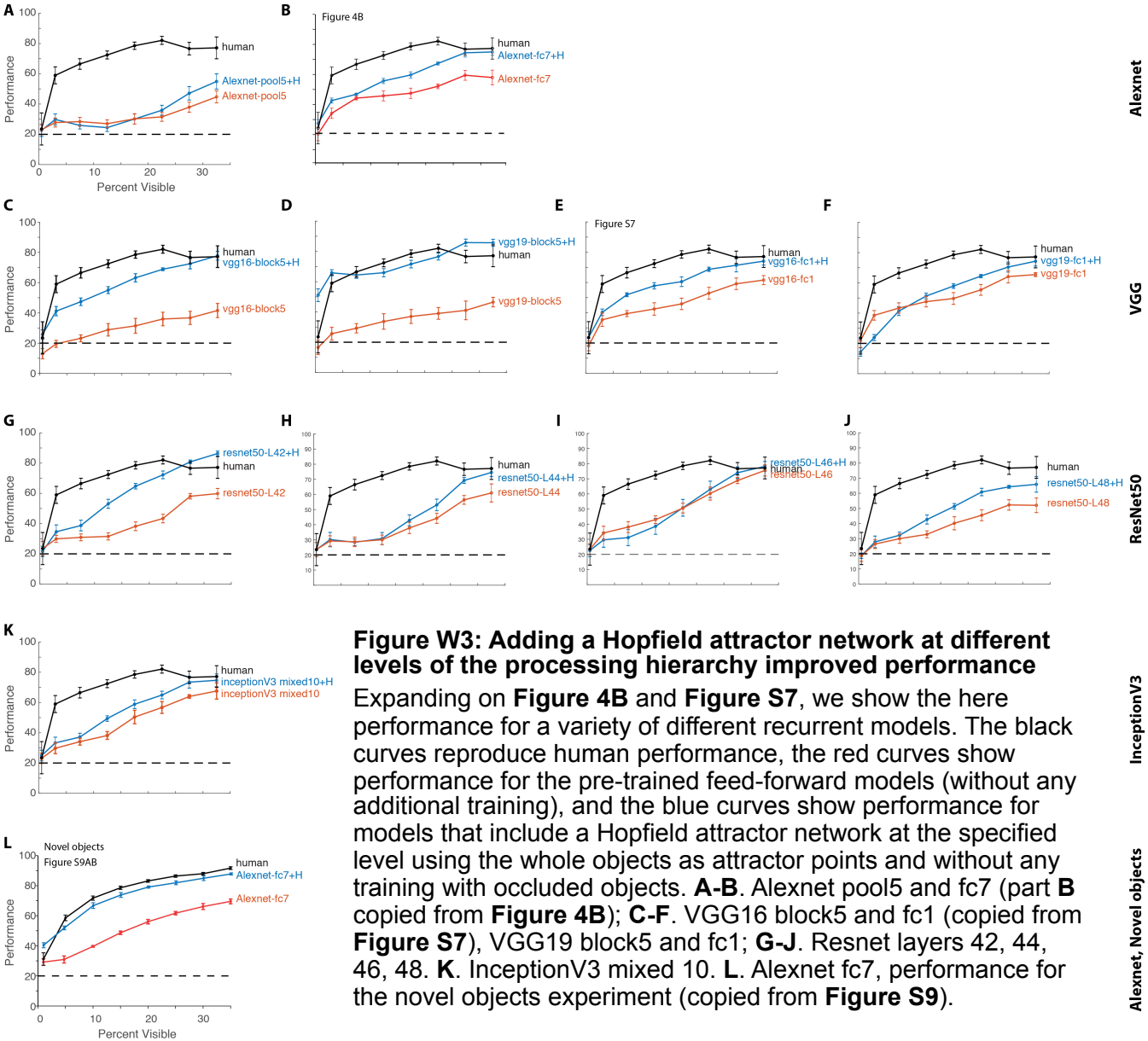

Figure W3: Adding a Hopfield attractor network at different levels of the processing hierarchy improved performance

Expanding on **Figure 4B** and **Figure S7**, we show the here performance for a variety of different recurrent models. The black curves reproduce human performance, the red curves show performance for the pre-trained feed-forward models (without any additional training), and the blue curves show performance for models that include a Hopfield attractor network at the specified level using the whole objects as attractor points and without any training with occluded objects. **A-B**. Alexnet pool5 and fc7 (part **B** copied from **Figure 4B**); **C-F**. VGG16 block5 and fc1 (copied from **Figure S7**), VGG19 block5 and fc1; **G-J**. Resnet layers 42, 44, 46, 48. **K**. InceptionV3 mixed 10. **L**. Alexnet fc7, performance for the novel objects experiment (copied from **Figure S9**).

Effect of binarization on performance

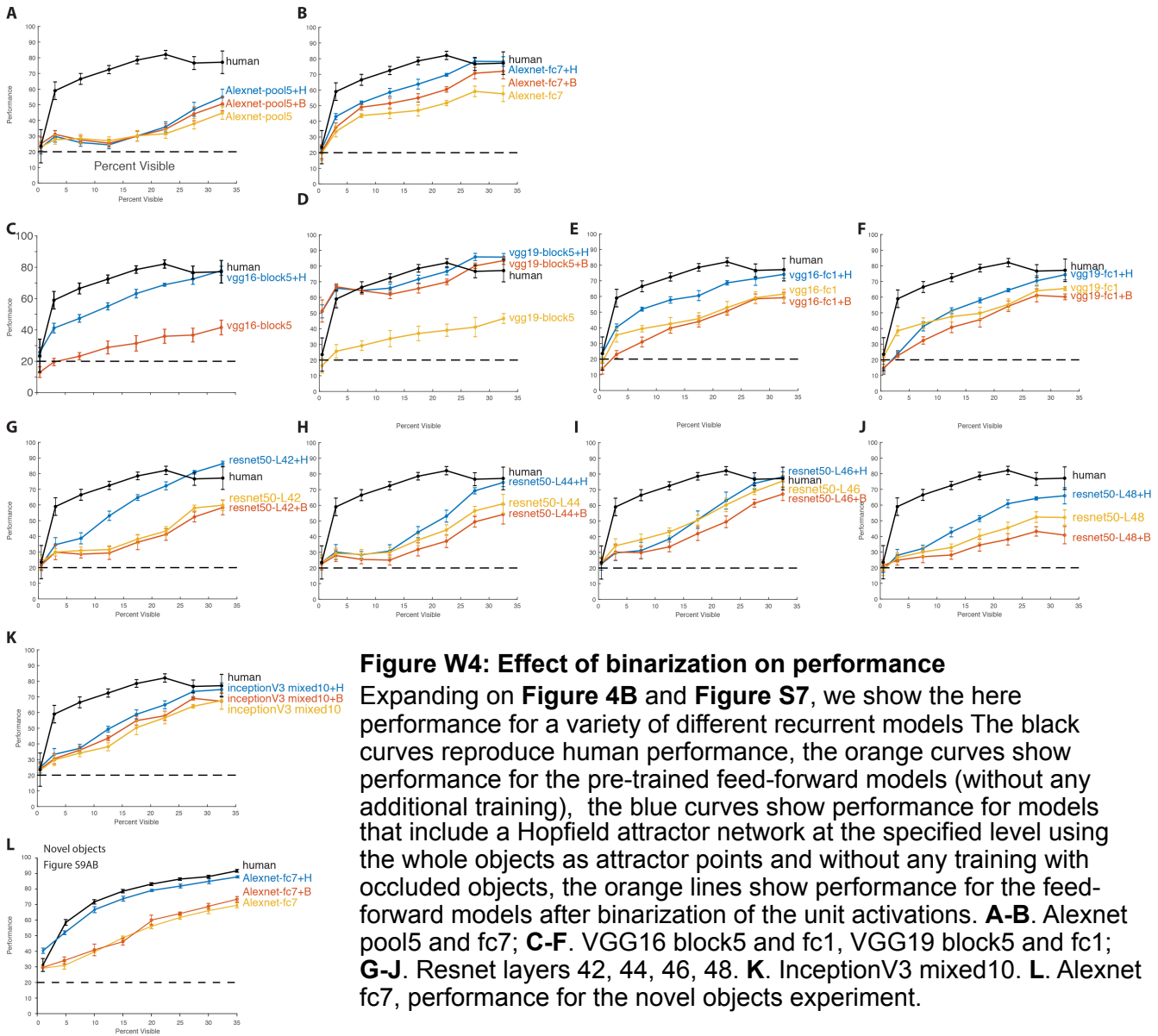

Figure W4: Effect of binarization on performance

Expanding on **Figure 4B** and **Figure S7**, we show the here performance for a variety of different recurrent models The black curves reproduce human performance, the orange curves show performance for the pre-trained feed-forward models (without any additional training), the blue curves show performance for models that include a Hopfield attractor network at the specified level using the whole objects as attractor points and without any training with occluded objects, the orange lines show performance for the feed-forward models after binarization of the unit activations. **A-B**. Alexnet pool5 and fc7; **C-F**. VGG16 block5 and fc1, VGG19 block5 and fc1; **G-J**. Resnet layers 42, 44, 46, 48. **K**. InceptionV3 mixed10. **L**. Alexnet fc7, performance for the novel objects experiment.

Effect of binarization on performance

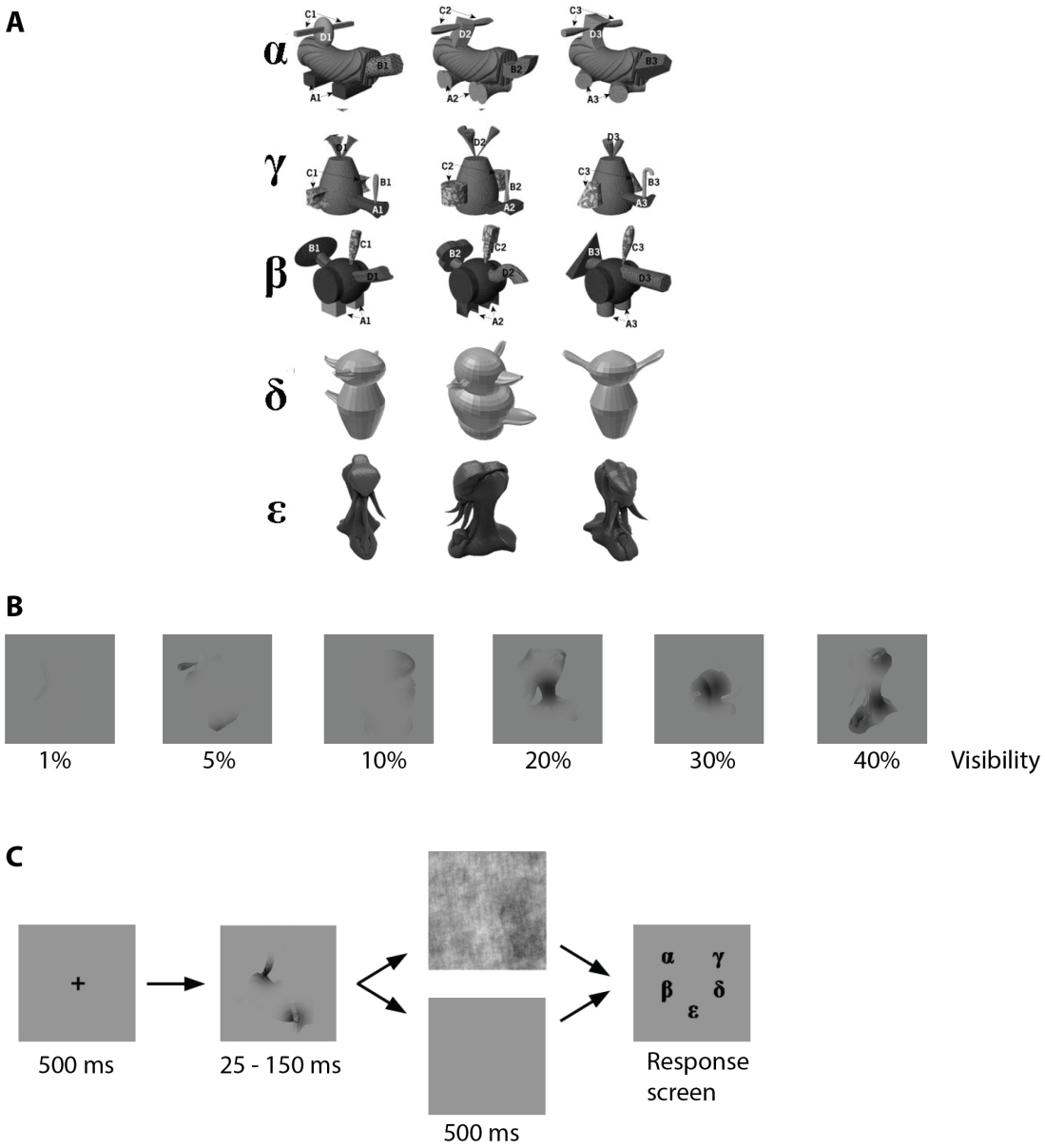

Figure W5: Examples of novel objects

A. Expanding on **Figure S8A**, 3 exemplar objects for each category of novel objects

B. Examples of novel objects rendered at different visibility levels.

C. Scheme of the psychophysics experiment using novel objects (similar to the structure shown in **Fig. 1A-B** for the main experiment).

Number of repetitions of novel objects before experiment

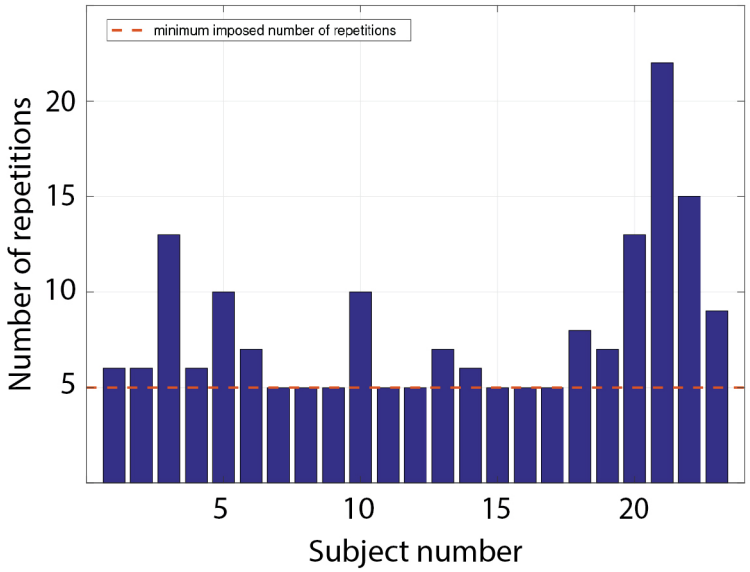

Figure W6: Subjects had only minimal exposure to 2 examples per category of whole novel objects

Number of repetitions per exemplar before the psychophysics test with partial novel objects started. Subjects were presented with 2 exemplars per category of whole novel objects to learn the mapping between categories and the 5 response buttons. In order to start the experiment, subjects were required to get 8 out of 10 correct responses, 5 times in a row using these practice stimuli. On average, reaching this level of accuracy required 80±40 trials.

Robust performance in recognition of partial novel objects was observed from the beginning

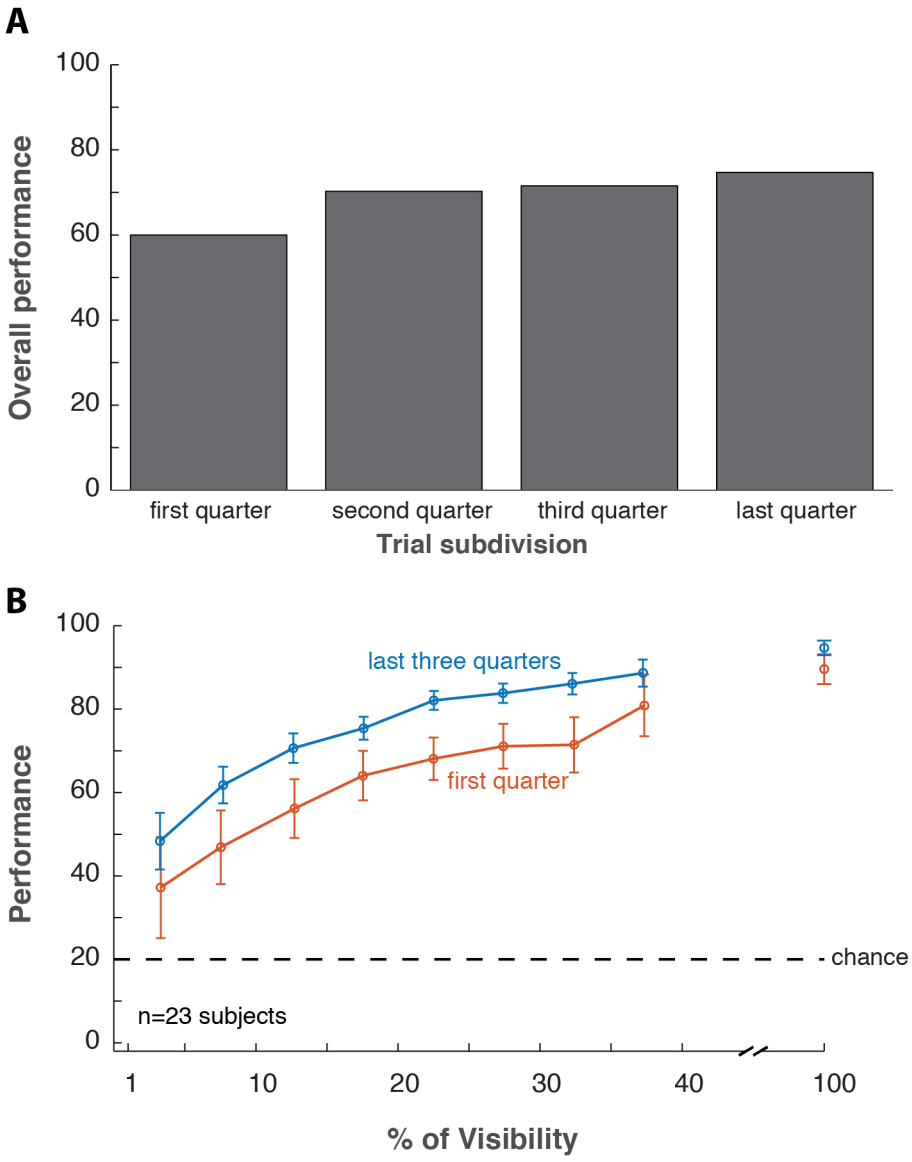

Figure W7: Robust performance from initial exposure to partial novel objects

We separated the psychophysics data into 4 quarters. Performance was well above chance levels across all visibility levels even in the first quarter, that is, even with the very first presentations of partial novel objects. Performance in the last 3 quarters was better than in the 1st quarter across all visibility levels. The enhanced performance in the last 3 quarters probably reflects general task improvement, familiarity with the task, and memory for the 5 keys that need to be mapped onto the 5 categories given that this improvement is also apparent for the whole novel objects.