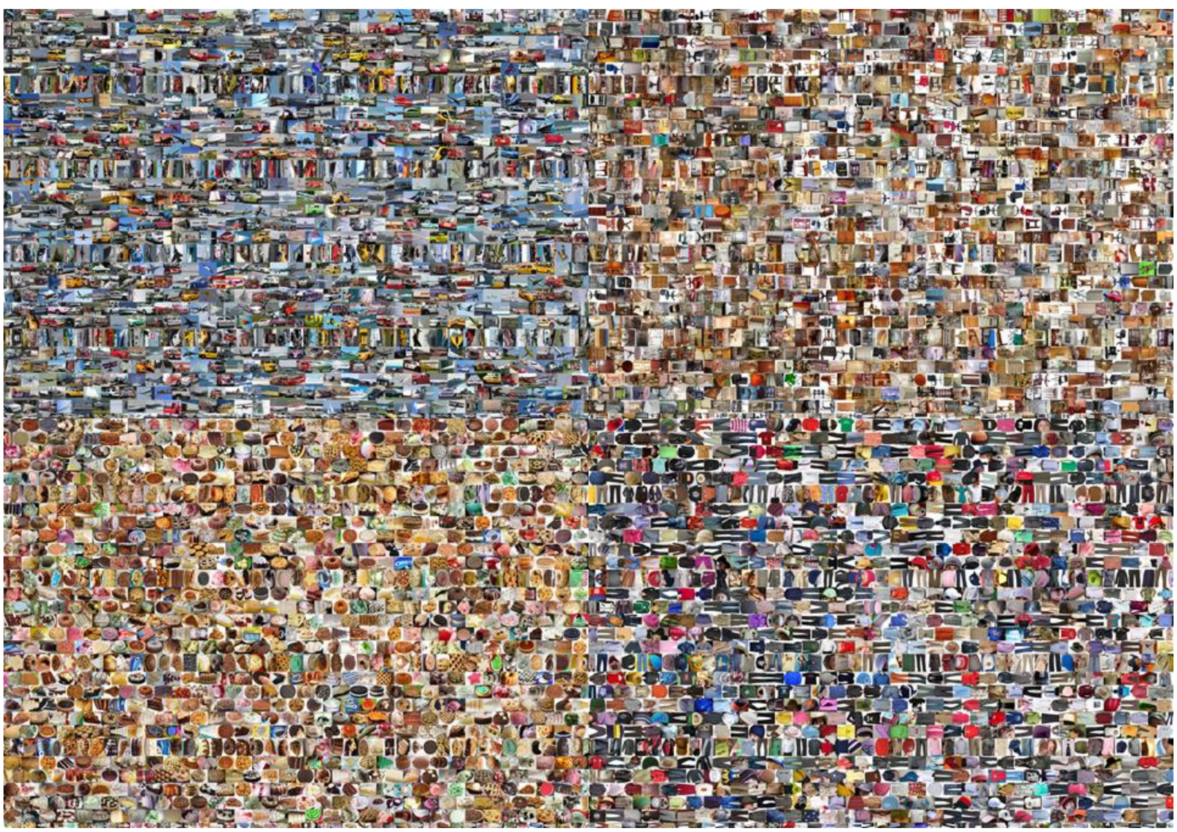
Supplemental Materials

A

B



C

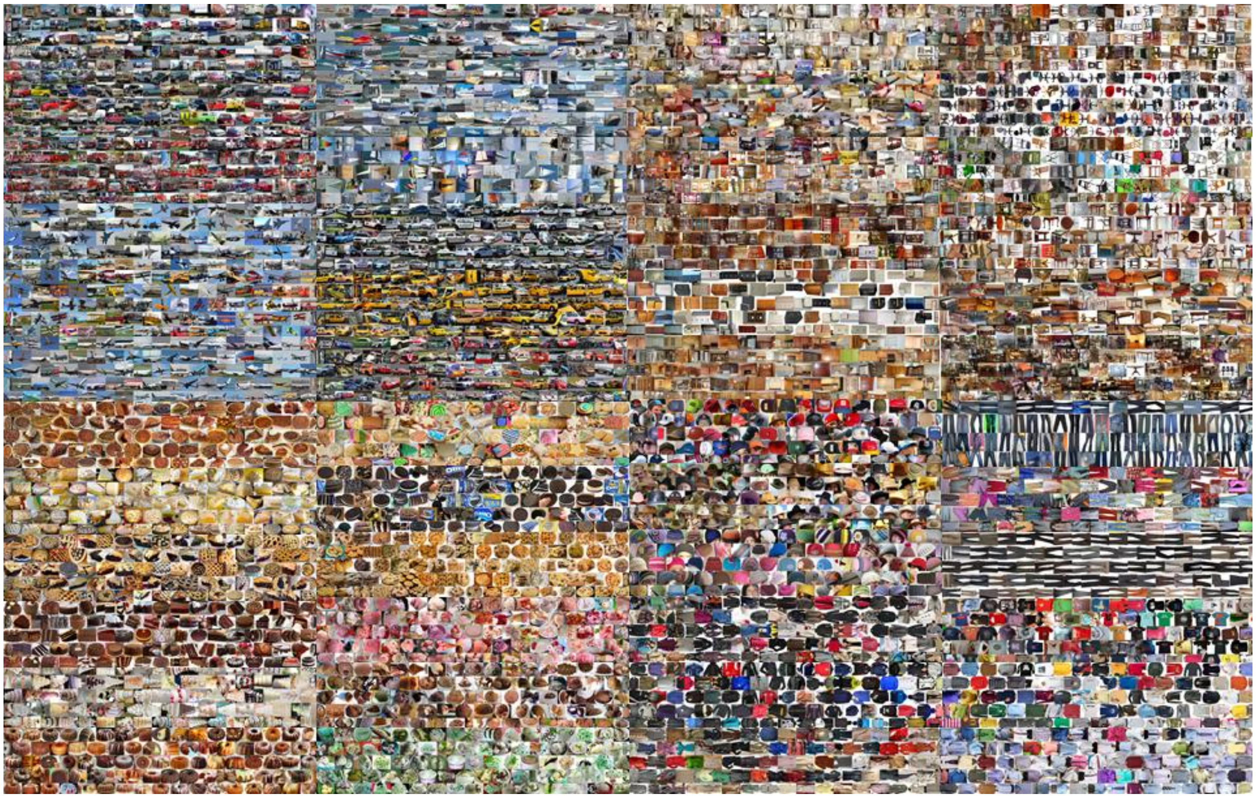


Figure S1. Tiny depictions of the ~4,800 images of objects used to train the CCF model. (A) Objects placed randomly. (B) Objects grouped into the 4 superordinate-level categories used as stimuli. (C) Objects grouped into the 48 subordinate-level categories used as stimuli. Note that in (B) between-category differences at the superordinate level create clearly separable visual regions, whereas in (C) some subordinate-level categories are visually distinct from siblings while many others are not.

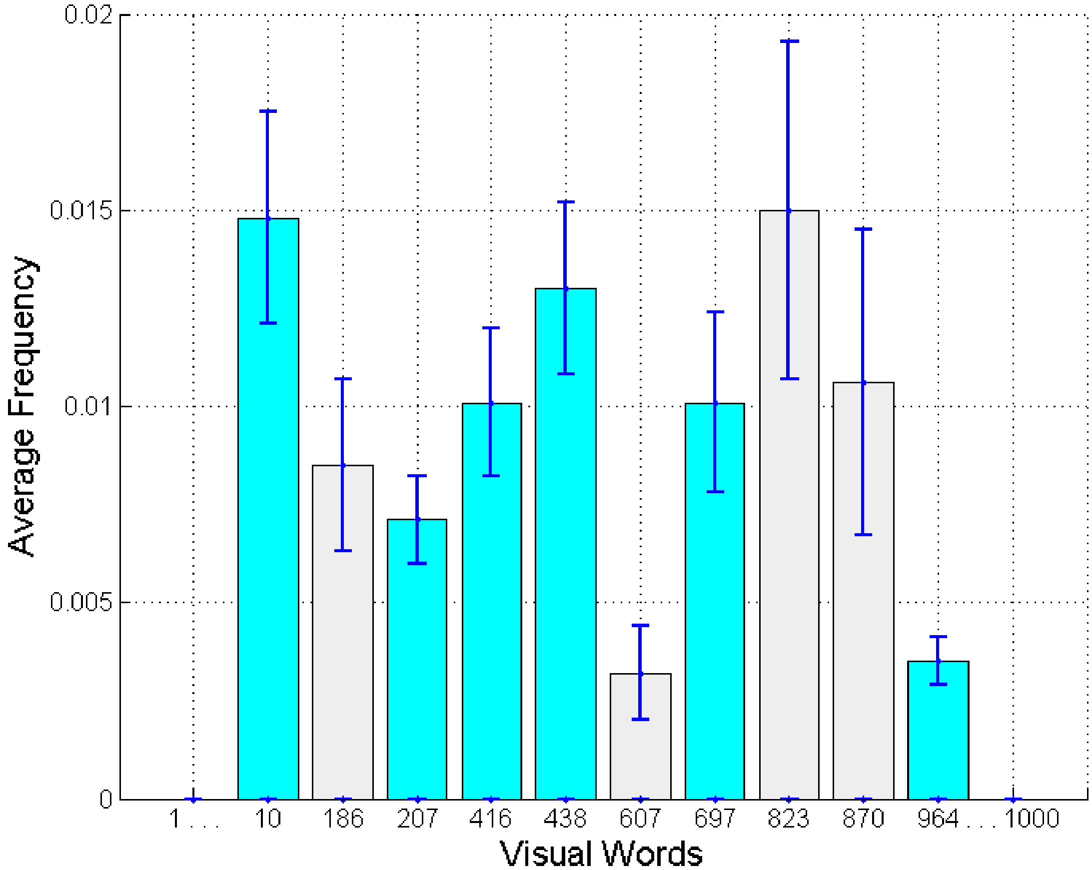


Figure S2. A partial plot (10 of the 1000 SIFT-feature bins) of the histogram obtained by averaging the 100 BoW histograms for the exemplars of the taxi category. Following this averaging each visual word bin has both a mean frequency and a variance (error bars indicate standard error). Blue bars indicate some of the visual words from the averaged histogram that would ultimately become CCFs for the taxi category representation. Gray bars indicate examples of visual words that were excluded from the CCF representation due to either excessive variability or too low of a mean frequency (see text for details). Note that frequencies are generally low because bin frequency sums to one for representation as a probability distribution. The average bin frequency is therefore 1/1000 or 0.001, considerably smaller than all of the bin frequencies shown.

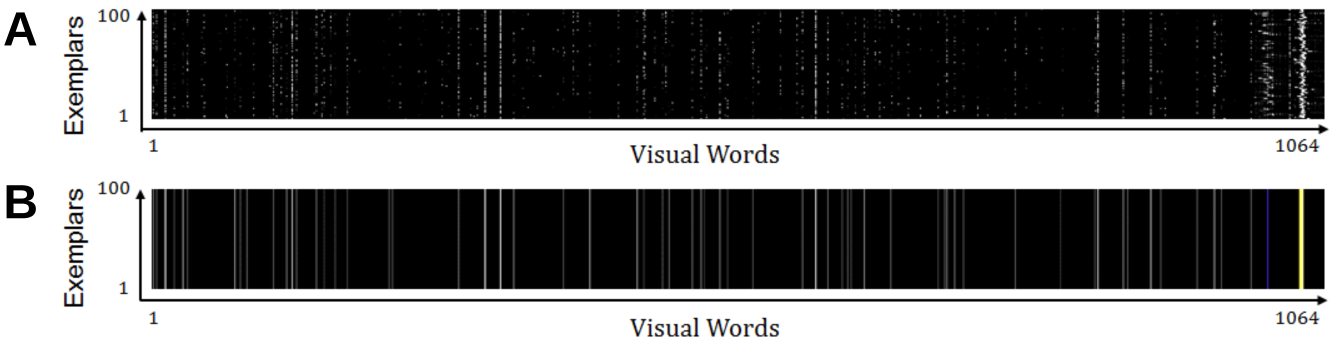


Figure S3. Histograms of all the exemplars from the taxi category shown "stacked" into a rectangular block. Each row shows the entire set of 1064 visual words corresponding to one exemplar, and each column shows the values for a given visual word for each of the 100 exemplars. (A) Averaged histograms of the type plotted in Figure S2 (blue and gray bins). Brighter pixels denote a higher frequency of features. Note that stacking the histograms makes it possible to visualize feature commonalities across exemplars, with features common to most exemplars appearing as bright vertical bars. Broken vertical bars indicate features appearing inconsistently across exemplars, and dark regions indicate features that are unimportant in the representation of the taxi category. (B) Stacked CCF histograms (only the blue bins from Figure S2). Vertical bars indicate the CCFs selected from the features in (A), with brightness indicating the mean frequency for a given visual word divided by its standard deviation. Bins 1001-1064 correspond to the hue visual words, where yellows appear prominently across the taxi exemplars. Note that the process of selecting CCFs serves to remove noise from the averaged histograms, thereby accentuating the features most important for the category representation.

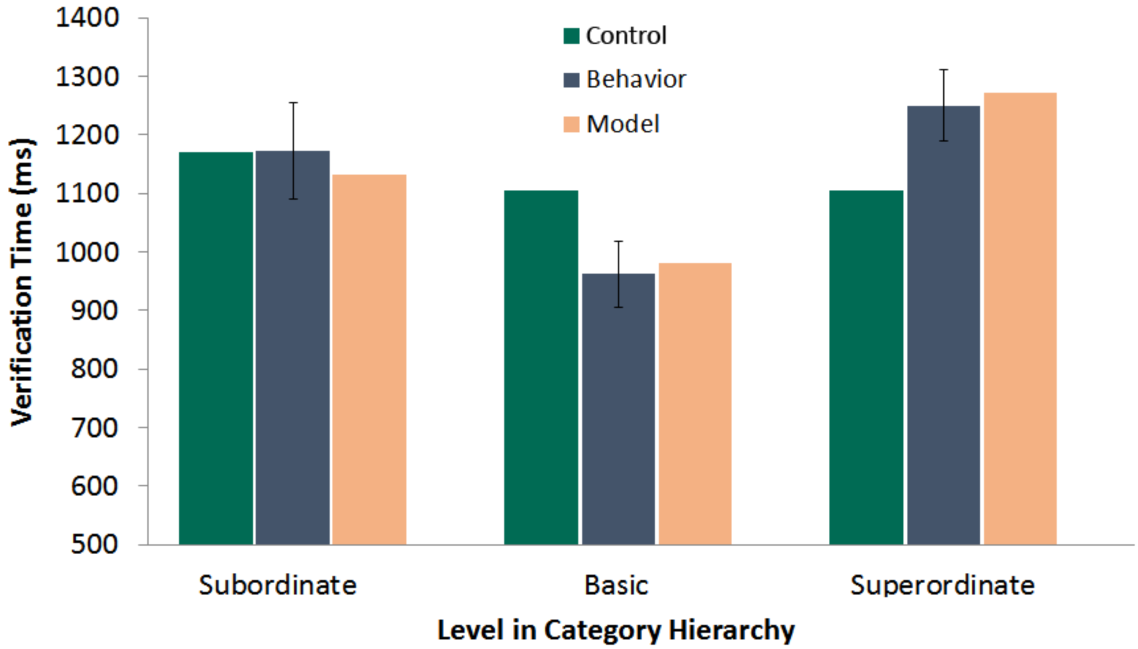


Figure S4. Results from a control experiment testing whether any features from the visual dictionary, regardless of whether they were CCFs, would be as good as those selected by the CCF model in characterizing the behavior data. The number of CCF features were found for each of the 68 categories, then the same number of features were randomly selected from the category's averaged BoW histogram (Figure S2 and S3A). These non-CCFs were used to compute sibling distances for each category, which were finally multiplied by the respective CCF number to get a category verification estimate, just as in the CCF model. Averaging these estimates by hierarchical level produced the data in green (leftmost bars), which show a much poorer fit to the behavioral data than the CCF model (re-plotted from Figure 3C). This experiment serves as a validation of the CCF approach; only CCFs were able to capture the BSE observed in behavior, the same numbers of randomly selected features could not.