What computational mechanisms underlie ensemble perception (Ariely, 2001)? All proposals to date rely on measurements of individuated objects, e.g., sizes of the circles in an ensemble, or their density in a region of space. These models of ensemble perception are also quite narrow, formulated in stimulus- and task-specific terms, such as averaging Gabor orientation or facial emotion. In light of mounting experimental evidence that human performance does not match object-based predictions (e.g., Cain, Dobkins, & Vul, VSS 2016), we seek to replace these models with one based on texture statistics (Portilla & Simoncelli, 2000; P&S). We recently built a texture statistics ideal observer (Cain & Cain, VSS 2018) that matches patterns of human successes and failures on the classic mean size comparison task (Chong & Treisman, 2003). Our ideal observer computes the P&S statistics for each ensemble, and uses the vector of their differences as features to classify which ensemble has larger mean size. With minimal training, a linear SVM is able to capture previously-reported biases in unequal setsize comparisons, providing the first principled account of these Point of Subjective Equality (PSE) shifts (Cain, et al., VSS 2016; see also Chong & Treisman, 2005; Sweeny, Wurnitsch, Gopnik, & Whitney, 2014).

Here we present a novel experiment that tests a key prediction of our model. If computation of texture statistics underlies ensemble perception, then random synthetic images constrained to match the texture statistics of the original ensembles should elicit similar responses from human observers as the original ensembles themselves. Indeed, naive participants were able to perform the larger mean circle size task on these "texsemble" stimuli directly, without ever seeing the ensembles of circles. Remarkably, we found good agreement of unequal setsize PSE shifts across human and ideal observers on ensemble and texsemble stimuli.

In sum, we are presenting the first image-computable model for ensemble perception, along with a strong behavioral test of the sufficiency of the statistical representation underlying it. Our single-pooling-region texture representation approach can account for a previously unexplained psychophysical effect, while providing support for texture-based mechanisms rather than rapid object processing. We will discuss implications for full-field texture models with many pooling regions (Freeman & Simoncelli, 2011), while placing these experiments in the theoretical context of the texture "default processing model" (Rosenholtz, 2014).