SPECIFIC AIMS

Object recognition is a magical biological feat: effortless and unexplained. A computational account for how the human brain recognizes objects would be an important milestone in neuroscience. However, it has proven to be a hard problem, that remains unsolved despite excellent physiological work by Connor, Desimone, DiCarlo, Gross, Perret, Poggio, and others. This proposal addresses a smaller problem, recognition of a simple object, like a letter. The Background section presents several lines of converging evidence for the idea that object recognition is always mediated by one or more highly stereotyped recognition units, and that a single unit is enough to recognize any simple object, like a letter, that lands in its receptive field. As noted in the Introduction to Revised Proposal, these recognition units all seem to be instances of a *computational kernel*, each applying the same computation to a different bit of the visual field. This proposal pursues the idea that peripheral recognition units are like foveal units, not worse, and much easier to isolate and study.

In crowding, a simple target (like a letter) that is easily recognized alone becomes impossible to recognize ("jumbled") when surrounded by clutter (anything similar), unless the clutter is a certain critical spacing away, which we call the *crowding distance*. This operationally defines the receptive field of the recognition unit that recognizes the target, as the area encompassed by the crowding distance in every direction. The recognition unit will recognize a single letter in its receptive field, but will not recognize anything if two letters or more are present. That's a severe information bottleneck in perception, and available evidence indicates that there is no bypass, other than using several units in parallel, each recognizing a part (e.g. a letter or nose), to recognize a complex object, like a word or face (Martelli et al., 2005). We define the "recognition unit" operationally as the computation underlying recognition of a simple target with and without clutter. The radius of its receptive field is the psychophysically measured crowding distance (deg).

In the periphery, recognition units are sparse and easily isolated, in practically every crowding study. Though dense in the fovea, we have managed to isolate there too, using the skinny new Pelli font (Aim 2). Our pilot data on conservation of efficiency and equivalent noise indicate that the recognition units perform the same computation regardless of location in the visual field (Aim 3). Our joint psychophysical-fMRI measurement of cortical crowding distance in mm, indicate that the recognition units have the same receptive field size (mm^2) on the surface of ventral occipital cortex (hV4, not V1-V3) independent of eccentricity and orientation (Aim 1). Furthermore, strabismic amblyopia and apperceptive agnosia impair object recognition by making central vision like normal peripheral vision, at some eccentricity. Thus, eccentricity and these two diseases affect only the sparseness of the recognition units.

This proposal is a full-scale assault on the recognition unit, coming at it from three directions, to characterize its neural substrate (Aim 1), relevance to public health (Aim 2), and computational model (Aim 3). All Aims are Years 1-3.

Aim 1. ***Atlas of Conservation of Cortical Crowding Distance***. In this Aim, we assess, in each visual area of the cortex, the degree of conservation of crowding distance (mm) across eccentricity, orientation, and participants. We measure crowding distance and acuity psychophysically as well as cortical magnification in 10 visual areas by fMRI, at 0°, ±4°, and ±8° horizontal eccentricity, radially and tangentially, in 10 healthy adults per year. The resulting atlas of the relationship between crowding distance and cortical magnification will constrain physiological models for crowding's source. Our pilot data indicate conservation of 2.6±1.5 mm (mean±sd) crowding distance across eccentricity and orientation in area hV4 (unlike V1, V2, and V3, which dramatically fail to conserve it across orientation). Across participants, we find a -0.5 correlation of log psychophysical crowding distance (deg) and log cortical magnification in hV4. Such conservation and correlation, if confirmed, would validate crowding distance (deg) as a quick, noninvasive measure of hV4 condition, clinically applicable during normal and abnormal development. In Years 2-3, we will also test severe cases of strabismic amblyopia. With Drs. Winawer (co-I) & Benson (NYU).

Aim 2. ***Correlating Crowding vs. Reading in normal development, amblyopia, dyslexia, & Posterior Cortical Atrophy (PCA).*** Reading a word requires one recognition unit per letter, so developmental and clinical changes in the crowding distance (the recognition unit's receptive field radius) could affect reading speed, especially in dyslexic children (Pelli et al. 2007; Pelli & Tillman, 2008). We will correlate crowding distance vs. reading speed for amblyopic, dyslexic, and normal children, ages 5 to 12, as well as for normal adults and those with PCA. Our new clinical test for crowding enables quick, accurate measurement of foveal crowding distance, with immunity to fixation errors (Pelli et al. 2016). The correlation of reading speed with crowding distance across diverse populations will expose the limits that crowding imposes on reading; yield norms for the development of crowding; and assess the degree to which strabismic amblyopia arrests the normal developmental reduction in crowding. Huge public interventions seek to help dyslexic children read faster and to identify amblyopic children sooner. Crowding can be measured *before* children learn to read. Our results will assess its potential to screen children for amblyopia and risk for dyslexia. With Drs. Waugh (Anglia Ruskin U., UK), Martelli (U. Rome, Italy), Rhodes (NYU), & Primativo (UCL, UK).

Aim 3. ***Psychophysics & Modeling. Conservation across the visual field and across size of receptive field.*** Following up our pilot data showing conservation across all eccentricities (evidence that the recognition unit is a computational kernel), we will measure efficiency and equivalent noise over 0° to 60° eccentricity, for sizes of 0.5° to 32° (as allowed by acuity), for letters and other simple objects on ten observers. In pilot data, we present a ConvNet machine-learning model of one human recognition unit, showing that it performs like human observers, to an uncanny degree. We continue testing this model of simple-object recognition, now focusing on the human's remarkable conservation of efficiency for target of given size by the human recognition units as receptive field size is hugely increased from fovea to periphery. With Dr. LeCun (co-I) & postdoc TBA.