Human vision is full of puzzles. Observers can grasp the essence of a scene in an instant, yet when probed for details they are at a loss. People have trouble finding their keys, yet they may be quite visible once found. How does one explain this combination of marvelous successes with quirky failures? I will describe our attempts to develop a unifying theory that brings a satisfying order to multiple phenomena.

The key is to understand peripheral vision. A visual system cannot process everything with full fidelity, and therefore must keep some information while losing other information. More than 99% of the visual field lies outside the fovea. Peripheral vision must condense this mass of information into a succinct representation that nonetheless carries the information that is needed for vision at a glance. My lab has proposed that the visual system deals with limited capacity by encoding the visual input in terms of a rich set of local image statistics, where the local regions grow — and the representation becomes less precise — with distance from fixation. This representation computes sophisticated image features at the expense of spatial localization of those features. This tradeoff is critical to modeling vision, and once understood, a great many visual phenomena can be explained without further ado. Furthermore, this encoding scheme has implications not only for the study of human vision, but also for applications such as multitasking with a mobile device, and design of effective user interfaces and information visualizations.