

How selective and non-selective pathways contribute to visual search in scenes.

Jeremy M. Wolfe, Melissa L.-H. Võ, Karla K. Evans
Brigham & Women's Hospital and Harvard Medical School

The world presents the visual system with far more information than the system can handle. Selective attention allows processing capacity to be directed to a subset of the input; for example, in order to identify an object. When we search for a specific target object among distractors, attention is “guided” to items that are likely to be targets on the basis of basic attributes of the stimuli such as color, size, and orientation. Thus, in a search for a **red vertical** target among **red** horizontal and green **vertical** distractors, attention can be guided to red items and to vertical items. The intersection of those two sets would be an excellent place to look for a red vertical item. This account is the heart of the “Guided Search (GS)” model that does a very credible job of accounting for a range of searches for targets presented in random arrays of distractors.

Random arrays of items on computer screens are not really what our visual “search engine” was built for. We usually search for objects in real-world scenes. The promise of GS was that it would scale-up to the complexity of the real world. There are, perhaps, 1 or 2 dozen guiding attributes. These define a high-dimensional space in which to locate objects. Search for a lion or a toaster would be a search for an N-dimensional conjunction of guiding features. Sadly, this proposal fails. Searches for arbitrary target objects in random arrays of distractor objects are too inefficient. Moreover, finding objects in scenes is easier than it should be. The structure of the scene is contributing to the guidance of attention in a manner not previously incorporated into GS.

In this talk, we will discuss two forms of “scene guidance”. Global spatial scene layout and scene statistics rapidly extracted across the scene by “non-selective” processes allow us to access stored knowledge about scenes that in turn we implicitly use to guide our search in scenes. One form of scene guidance is “Semantic” guidance – guidance based on meaning: E.g., forks are often found next to knives. The other is “Syntactic” guidance – guidance based on the structure/physics of the scene: E.g. forks do not float in mid-air. They require support. The result of this analysis is a two-pathway version of GS that incorporates selective and non-selective pathways to visual awareness.

