Commentary on Hulleman & Olivers "*The impending demise of the item in visual search*"

# What fixations reveal about oculomotor scanning behavior in visual search

Zhuanghua Shi[1], Xuelian Zang[1,2] & Thomas Geyer[1]

1. General and Experimental Psychology, Department of Psychology, LMU Munich
2. China Centre for Special Economic Zone Research, Shenzhen University, China

## Abstract

Hulleman and Olivers' (H&O) conceptual framework does not consider variation of fixation duration and its interaction with the size of the functional viewing field (FVF). Here we provide empirical evidence of a dynamic interaction between the two parameters, suggesting that fixations, as the central unit in H&O's framework, should be studied on both the spatial and temporal dimensions.

## Commentary

By taking fixations, not individual items, as the central unit, Hulleman & Olivers (H&O) put forward a promising, unified account of both eye-movement behavior and manual reaction times (RTs) in visual search. However, their conceptual framework makes two simplified assumptions: (1) the size of the functional viewing field (FVF) being solely dependent on the visual discriminability of the search elements; and (2) constant FVF processing time (i.e., a constant fixation duration of 250 ms), ignoring any dynamic interactions between the two parameters. Although the assumption of constancy of fixation durations makes the framework easily comparable with traditional, item-based selection models, it limits the explanatory potential of H&O's account – as we will outline in this commentary.

It is generally accepted that 'fixate' and 'move' oculomotor activities are governed by parallel 'when' and 'where' commands generated across the entire visual-perceptual hierarchy (Findlay & Walker, 1999). Concerning top-down influences, fixation durations are influenced by task difficulty (Hooge & Erkelens, 1998; Moffitt, 1980; Pomplun, Garaas, & Carrasco, 2013), memory about spatial context (van Asselen, Sampaio, Pina, & Castelo-Branco, 2011; Zang, Jia, Müller, & Shi, 2015), visual search strategy (Geyer, Von Mühlenen, & Müller, 2007), and multisensory experience (Zou, Müller, & Shi, 2012). For example, Geyer et al. (2007) compared fixation durations between static and dynamic search displays with identical target-distractor discriminability, except that search items were randomly reshuffled every 117 ms in the latter condition. Mean fixation duration, as well as the latency of the first saccade, was increased by some 100-150 ms for the dynamic compared to the static condition, although 'standard' measures of search efficiency (slope of the search function) were comparable between the two types of display. These findings clearly suggest that fixational dwell times are not solely under the control of the current sensory environment, or in H&O's terms: the perceptual discriminability of the search items. Instead, observers' strategic efforts in solving the task at hand must be considered, too, in accounting for such extended fixation durations (Geyer et al., 2007).

In most cases, fixation duration and the FVF size, rather than being independent, interact in a nonlinear fashion (Nuthmann, Smith, Engbert, & Henderson, 2010; Unema, Pannasch, Joos, & Velichkovsky, 2005). One strong piece of evidence of a dynamic interaction between the two parameters comes from an oculomotor study on the ‘pip-and-pop’ effect (Zou et al., 2012). In ‘pip-and-pop’ visual search displays, beeps are synchronized with (task-irrelevant) color changes of the target, which is presented in a cluttered and heterogeneous item field (with search being extremely ‘inefficient’). Zou et al. found fixation durations to be increased, by some 150 ms, for beep-present vs. beep-absent trials: an ‘oculomotor freezing’ effect. Such extended fixations at beeps allow information to be sampled over a larger region (i.e., increasing the FVF), as indicated by larger saccade amplitudes immediately after the beeps. In other words, beep-induced prolonged fixation times and subsequent large saccade amplitudes mediate fast detection of target presence, yielding the ‘pip-and-pop’ effect. This pattern also suggests that the oculomotor scanning strategy can affect the rate of information processing, as evidenced by increased information up-take per fixation for the beep-present relative to the beep-absent condition. Another very recent study (Zang et al., 2015) on context-based guidance of visual search also revealed a beneficial effect of extended fixation duration on task performance. In this study, observers were first trained with an artificial FVF size, implemented by a gaze-contingent tunnel-viewing technique. With 4-5 items visible inside the FVF, the mean fixation duration was extended already in the training session for repeated ‘old’, compared to randomly generated ‘new’, display (item) layouts. Further, the scanpath for old relative to new displays was closer to the optimal scanpath, indicating that learned context improves the efficiency of oculomotor scanning. Increased fixational dwell times and shortened scanpaths for old relative to new displays remained evident even after the constraining tunnel view was removed from the task. – Such dynamic adjustments of fixation duration and saccade amplitude are quite common during scene search. It has been shown, for instance, that fixation duration and saccade amplitude gradually change over the first few seconds, and then approach their asymptotic levels (Unema et al., 2005). Both asymptotes, however, depend on the number of objects in the scene – which indicates that the complexity of the scene, too, changes oculomotor scanning.

These findings, amongst others, provide converging evidence that the size of FVF and fixation duration are not determined by visual discriminability alone, as assumed by H&O. Rather, oculomotor scanning behavior is dynamic in that the size of FVF and fixation duration must be considered together to discern moment-by-moment adjustments of information processing. Despite H&O's conceptual framework (as yet) lacking flexibility of oculomotor parameters, the idea of fixation as a central processing unit of visual search remains very promising. However, in order to incorporate the above findings of dynamic interactions between fixation duration and saccade amplitude, we propose that fixational eye movements are best characterized by both spatial (i.e., the size of FVF in H&O terms) and temporal (i.e., fixation duration) factors. Combining the two could provide insight into how oculomotor scanning strategies influence the fixation-by-fixation information processing rate, which might turn out to be the distinguishing feature for comparing different visual search tasks. (words: 866)

**References**


Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, *22*(04), 661–674.

Geyer, T., Von Mühlenen, A., & Müller, H. J. (2007). What do eye movements reveal about the role of memory in visual search? *Quarterly Journal of Experimental Psychology*, *60*(7), 924–935.

Hooge, I. T. C., & Erkelens, C. J. (1998). Adjustment of fixation duration in visual search. *Vision Research*, *38*(9), 1295–1302. doi:10.1016/S0042-6989(97)00287-3

Moffitt, K. (1980). Evaluation of the fixation duration in visual search. *Perception & Psychophysics*, *27*(4), 370–372. doi:10.3758/BF03206127

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review*, *117*(2), 382–405. doi:10.1037/a0018924

Pomplun, M., Garaas, T. W., & Carrasco, M. (2013). The effects of task difficulty on visual search strategy in virtual 3D displays. *Journal of Vision*, *13*(2013), 1–22. doi:10.1167/13.3.24.doi

Unema, P. J. a., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, *12*(3), 473–494. doi:10.1080/13506280444000409

van Asselen, M., Sampaio, J., Pina, A., & Castelo-Branco, M. (2011). Object based implicit contextual learning: a study of eye movements. *Attention, Perception & Psychophysics*, *73*(2), 297–302. doi:10.3758/s13414-010-0047-9

Zang, X., Jia, L., Müller, H. J., & Shi, Z. (2015). Invariant spatial context is learned but not retrieved in gaze-contingent limited-viewing search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 807–819. doi:10.1037/xlm0000060

Zou, H., Müller, H. J., & Shi, Z. (2012). Non-spatial sounds regulate eye movements and enhance visual search. *Journal of Vision*, *12*(5), 2, 1–18. doi:10.1167/12.5.2