Acta Psychologica 157 (2015) 131–143

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

You can detect the trees as well as the forest when adding the leaves: Evidence from visual search tasks containing three-level hierarchical stimuli

Claire-Sara Krakowski [a], Grégoire Borst [a], Arlette Pineau [a], Olivier Houdé [a,b], Nicolas Poirel [a,b,*]

[a] LaPsyDÉ, UMR 8240, CNRS, Université Paris Descartes, Université de Caen, France
[b] Institut Universitaire de France (IUF), Paris, France

ARTICLE INFO

Article history:
Received 23 April 2014
Received in revised form 16 February 2015
Accepted 4 March 2015
Available online 18 March 2015

PsychINFO codes:
2323
2346

Keywords:
Global processing
Local processing
Intermediate processing
Visual search task

ABSTRACT

The present study investigated how multiple levels of hierarchical stimuli (i.e., global, intermediate and local) are processed during a visual search task. Healthy adults participated in a visual search task in which a target was either present or not at one of the three levels of hierarchical stimuli (global geometrical form made by intermediate forms themselves constituted by local forms). By varying the number of distractors, the results showed that targets presented at global and intermediate levels were detected efficiently (i.e., the detection times did not vary with the number of distractors) whereas local targets were processed less efficiently (i.e., the detection times increased with the number of distractors). Additional experiments confirmed that these results were not due to the size of the target elements or to the spatial proximity among the structural levels. Taken together, these results show that the most local level is always processed less efficiently, suggesting that it is disadvantaged during the competition for attentional resources compared to higher structural levels. The present study thus supports the view that the processing occurring in visual search acts dichotomously rather than continuously. Given that pure structuralist and pure space-based models of attention cannot account for the pattern of our findings, we discuss the implication for perception, attentional selection and executive control of target position on hierarchical stimuli.

1. Introduction

In every-day life, we are continuously confronted with complex visual scenes. To act efficiently within our environment and achieve our goals, we must select relevant information in those complex visual scenes (Eckstein, 2011). Indeed, imagine you have to find pertinent information on a website page full of advertisements, texts, and pictures of different sizes that are sometimes fitted into each other. In spite of the numerous distractors presented on the screen, we are particularly efficient at finding and selecting the information that we are seeking. A critical question, then, is how do we achieve this fast and efficient detection of pertinent information?

Over the last century, psychologists have tried to understand how attentional resources are distributed in a visual display. Some of the previous studies have historically adopted a structuralist view, arguing that traits are registered before an object can be recognized. Treisman and Gelade (1980) used visual search tasks in which participants had to press a button if they detected a specific target and another button if the target was absent from the display. The target appeared in the display among a various number of visual distractors. The study found that when a target only differed from the distractors with respect to one simple feature (e.g., color, shape or orientation), the time required to detect this target was not dependent on the number of distractors presented in the display. Conversely, when a target differed from distractors by a conjunction of several features, the time required for detection increased with the number of distractors. They concluded, in their well-known structuralist "feature integration theory" (FIT), that features are registered in a parallel manner of processing. If the target differs from the distractors by only one simple feature, the target pops-out automatically (i.e., parallel processing), but if the target differs with respect to several features, attention is required to bind the features in a specific location, and thus, identify the whole object that is formed by them (i.e., serial processing). According to this view, attentional selection is space-based, an assumption shared by many other classical theories of attention (for example, LaBerge & Brown, 1989; Posner, 1980). Even though the FIT explained a multitude of behavioral results and perceptive phenomena, Wolfe, Cave, and Franzel (1989) showed that a target could pop-out from the distractors even though it differed based on a combination of several features. The salience of the target from the distractors seems to impact the visual search more than the number of features by which they differ (Desimone & Duncan, 1995; Wolfe, 2007).

* Corresponding author at: Laboratoire A. Binet, Sorbonne, 46 rue Saint Jacques, 75005 Paris, France.
E-mail address: nicolas.poirel@parisdescartes.fr (N. Poirel).

http://dx.doi.org/10.1016/j.actpsy.2015.03.001
0001-6918/

Several models also conceived visual search beyond the serial/parallel dichotomy (see Wolfe, 2007) because of the difficulty presenting strong evidence that a search slope could only represent the signature of a serial or a parallel processing. For example, a 5 ms by item search slope is often considered a parallel processing, but it could be interpreted as a very fast serial processing (Townsend, 1990). Moreover, search slopes that increase with the number of distractors could be interpreted as the result of limited parallel processing capacity (as opposed to serial processing) (see Bundesen, Habekost, & Kyllingsbæk, 2005; Thornton & Gilden, 2007) or as the result of a lower discriminability between target and distractors (which require attention to restrict the range of units responding to the distractors; see signal detection theories, Eckstein, 2011 and Verghese, 2001). Moreover, there is no evidence that linear search slopes represent purely serial processing; it could be a mix of parallel search and serial search (Townsend, 1990). Serial search could also be a group-by-group item search rather than an item-by-item search. Treisman (1993) suggested that an item could be processed in parallel inside a group of stimuli but that this search could be serial between two groups of stimuli. Accordingly, eye movements during visual search tasks indicated that parallel processing occurs inside a fixation and that serial processing is due to saccades (Young & Hulleman, 2013). The authors explained that feature search and conjunctive search differ because the number of objects processed during a fixation varies with task difficulty. Considering these new empirical data and models of visual search, various authors (see Eckstein, 2011; Wolfe, 2007) use the terminology "efficient search" if search time does not increase with distractors and "less efficient search" if search time increases with the number of distractors. This terminology is therefore employed in the present study.

Several studies also demonstrated that the consideration of the whole global object impacts the processing of features during the very first processing stages in opposition to what is expected by the FIT. For example, a complex feature such as the direction of lighting in the items could pop-out if the global form is three-dimensional, whereas a less efficient search is observed if the global form is two-dimensional (Enns & Rensink, 1990). These results with search tasks are in line with a study using a foveal presentation that demonstrated a faster processing for two properties when they are grouped within the same object than when the two properties appear in two different objects, although they overlap (Duncan, 1984). By providing evidence of an impact of the object itself on the processing of its features, this amount of data demonstrated a limitation of a pure structuralist approach toward visual search processes and space-based theories of attentional selection. A consensus now appears to be emerging, considering that attention operates upon spatial properties as well as organizational structures (see Yeari & Goldsmith, 2011).

Grouping effects reported in visual search tasks (see Treisman, 1993) are consistent with the effects reported in studies on global and local processing of visual information. Navon (1977), for example, found that with hierarchical stimuli (e.g., a large global "H" made of small local "S"s) a global form is detected more easily and more quickly than the local elements that compose it. According to Navon (1977) and others (see, for instance Kimchi, 1992; Poirel, Pineau, & Mellet, 2008), this global precedence effect (GPE) reflects two different aspects of local/global processing: (a) that a global target is detected more quickly than a local target — i.e., the so-called global advantage; and (b) that participants process the local target more slowly if conflicting information is simultaneously presented at the global level — i.e., the so-called global interference. These two aspects of the GPE depend on different types of processing (see Poirel, Mellet, Houdé, & Pineau, 2008), which is consistent with the model proposed by Bullier (2001) on monkeys' brains and recently applied to human brains (Peyrin et al., 2010). The global advantage seems to be directly related to the functioning of our visual system. Indeed, low-spatial frequencies that convey global information are processed faster than high-spatial frequencies that convey local information (see Hughes, Nozawa, & Kitterle, 1996).

As global information is processed faster than local information, a first sketch of the recognition of the visual scene and the preparation of a behavioral response are made on the basis of the global information. When the local information is subsequently processed, this information can either confirm or contradict the preliminary recognition of the scene based on global processing. At this stage, if global and local information do not match, global interference occurs. Therefore, global interference (as opposed to global advantage) seems related to top-down processes (see Beaucousin et al., 2011, 2013), which could include executive processes that enable the processing of local information when global information interferes. This hypothesis is supported by the results of a negative priming study suggesting that the global object must be inhibited to select objects at the local level (Poirel et al., 2014). Interestingly, only intensive inhibition training (more than 10,000 training trials) for global level information can reduce global interference (Dulaney & Marks, 2007). These behavioral data are supported by recent neuroimaging studies: when a participant has to focus her/his attention on a local letter amidst global interference compared with a situation in which the global letter is congruent, a very early change of amplitude on the N1 component is observed (Beaucousin et al., 2013) and brain regions implicated in conflict become more activated (Weissman, Giesbrecht, Song, Mangun, & Woldorff, 2003).

Executive processes also play a critical role in visual search tasks, particularly in regard to ignoring distractors to be able to correctly select a target. In the new FIT, inhibition is the mechanism by which the object file is updated (Treisman, 1993). The results of a negative priming task confirmed this hypothesis (DeSchepper & Treisman, 1996); specifically, participants required more time to compare a target to an object that they had previously just ignored. Thus, the selection of a relevant object is preceded by the inhibition of interfering ones, at least when these objects are equally (or more) salient. This behavioral data received support from electrophysiological (cells that process distractors are inhibited when detecting the target, see Desimone & Duncan, 1995) and fMRI (activations of the selective object cortex decrease when the corresponding object has to be ignored, see Seidl, Peelen, & Kastner, 2012) studies.

Executive control appears sufficiently implicated in visual search and attentional tasks that Desimone and Duncan (1995) proposed a biased competition model in which this process has a major role. These authors argued that all stimuli in the visual field are in competition due to a limited capacity for processing information from initial sensory inputs to motor responses. Several weights determined by bottom-up processes (that constrain visual perception) and by top-down processes, are attributed to the stimuli in this competition. When seeking an object within a complex visual environment, working memory holds its representation in a template, and visual competition is thus favored for this object. If a target looks like the distractors, all these pieces of information fit into the template, and detecting the target becomes more difficult. In this model, attention is not considered as a process that acts at a given time but rather as a slow process that manages the competition between elements. In this perspective, attention is closely related to the concept of executive control as described by Diamond (2013).

Additionally, the results obtained from the visual search and global–local detection tasks clearly demonstrate that the global object is not processed in terms of the local details it contains. To date, the exploration of how a target is detected depending on its level occurrence in a complex environment has not been well documented.

In a feature visual search task (i.e., a task in which the target differs from distractors by only one feature), Kimchi, Hadad, Behrmann, and Palmer (2005) showed that, when the global form contained many small elements (a 'good form' according to the gestalt laws of perceptual organization, see Wertheimer, 1938) global processing was effortless and efficient, whereas local processing was effortful and less efficient. The linear exploration observed during local processing suggests that local distractors are processed and ignored one by one (or group by group) until the local target is found, whereas the more salient global