Cognition 180 (2018) 91–98

Contents lists available at ScienceDirect

Cognition

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

Original Articles

Cognitive development attenuates audiovisual distraction and promotes the selection of task-relevant perceptual saliency during visual search on complex scenes

Clarissa Cavallina[a], Giovanna Puccio[a], Michele Capurso[a], Andrew J. Bremner[b,c], Valerio Santangelo[a,d,*]

[a] Department of Philosophy, Social Sciences & Education, University of Perugia, Italy
[b] Department of Psychology, Goldsmiths, University of London, UK
[c] School of Psychology, University of Birmingham, UK
[d] Neuroimaging Laboratory, Santa Lucia Foundation, IRCCS, Rome, Italy

ARTICLE INFO

Keywords:
Visual search
Complex scenes
Crossmodal spatial attention
Perceptual saliency
Cognitive development

ABSTRACT

Searching for a target while avoiding distraction is a core function of selective attention involving both voluntary and reflexive mechanisms. Here, for the first time, we investigated the development of the interplay between voluntary and reflexive mechanisms of selective attention from childhood to early adulthood. We asked 6-, 10-, and 20-year-old participants to search for a target presented in one hemifield of a complex scene, preceded by a task-irrelevant auditory cue on either the target side (valid), the opposite side (invalid), or both sides (neutral). For each scene we computed the number of salient locations (NSL) and the target saliency (TgS). All age groups showed comparable orienting effects ("valid minus neutral" trials), indicating a similar capture of spatial attention by valid cues which was independent of age. However, only adults demonstrated a suppression of the reorienting effect ("invalid minus neutral" trials), indicating late developments in the reallocation of spatial attention toward a target following auditory distraction. The searching performance of the children (both 6- and 10-year-olds), but not of the adults, was predicted by the NSL, indicating an attraction of processing resources to salient but task-irrelevant locations in childhood; conversely, only adults showed greater performance with increased TgS in valid trials, indicating late development in the use of task-related saliency. These findings highlight qualitatively different mechanisms of selective attention operating at different ages, demonstrating important developmental changes in the interplay between voluntary and reflexive mechanisms of selective attention during visual search in complex scenes.

1. Introduction

Selective attention, the ability to search for a target while ignoring distracting stimuli, is a core function of human cognition. In the last decades, a significant body of research has been devoted to understanding the role of selective attention in visual search (Eimer, 2014). There is now a wide consensus that selective attention can affect visual search either voluntarily, i.e., when attention is driven by visual search goals, or involuntarily (reflexively), i.e., when attention is driven by the physical saliency of the stimuli (e.g., Anderson, 2013; see also Klein & Shore, 2000). During search, there is a tight interplay between these two components of selective attention. In fact, voluntary attention devoted to search for a current target (or targets) is continuously challenged by distractors that can involuntarily capture available processing resources.

There is a clear consensus that involuntary aspects of selective attention are relatively stable across childhood and adulthood with voluntary aspects following a U-shaped pattern of development (improvements to young adulthood and decline into old age; e.g., Enns, Brodeur, & Trick, 1998; Rueda et al., 2004; Waszak, Li, & Hommel, 2010). In the domain of visual search, this issue has been typically assessed by comparing search for conjunctions of visual features vs. single features. While the former has been thought to provide an index of top-down attentional control, the latter has been considered to be mediated primarily by bottom-up involuntary orienting. For instance, Donnelly, Cave, Greenway, Hadwin, & Sonuga-Barke (2007) compared feature conjunction visual search performance between child and young adult participants. When searching for conjunctions of features

* Corresponding author at: Dept. of Philosophy, Social Sciences & Education, University of Perugia, Piazza G. Ermini, 1, 06123 Perugia, Italy.
E-mail address: valerio.santangelo@unipg.it (V. Santangelo).

https://doi.org/10.1016/j.cognition.2018.07.003
Received 21 September 2017; Received in revised form 11 June 2018; Accepted 4 July 2018

the children searched more slowly and took longer to reject trials than adults did when no target was present, indicating an age-related development of top-down control in visual search. By contrast however, searching for single visual features shows very little, if any development between childhood and adulthood (Trick & Enns, 1998). This pattern of findings is consistent with the general notion that voluntary visual search performance improves with age as a consequence of increasing attentional control across development (for reviews see Atkinson & Braddick, 2012; Scerif, 2010). While this important literature has clarified the distinct developmental trajectories of voluntary and involuntary attention, the development of the interaction of these two modes of attentional selection during visual search has been largely neglected to date.

One means of examining the interplay between voluntary and involuntary selective attention during visual search is by manipulating the deployment of visual spatial attention through a cuing procedure. In adults, this has typically been assessed by presenting either visual (Briand & Klein, 1987; Carrasco & Yeshurun, 1998; Cohen & Ivry, 1989; Müller & Rabbitt, 1989; Prinzmetal, Presti & Posner, 1986) or auditory (Perrott, Saberi, Brown & Strybel, 1990; Rudmann & Strybel, 1999; Vu, Strybel & Proctor, 2006) spatial cues before the presentation of the visual search scene. For instance, in their seminal study Prinzmetal et al. (1986) asked participants to search for a colored letter among distractor letters. On most trials, the participants were pre-cued by a visual peripheral signal to the spatial location were the target letter was presented (valid trial), while on the remaining trials a non-target location was cued (invalid trial). Prinzmetal and colleagues found more accuracy in valid than invalid trials. Although clearly demonstrating an effect of selective spatial attention on visual search, this study failed to demonstrate whether the effect was driven by: (i) a facilitation of visual search by selective attention, (ii) distraction by task-irrelevant stimuli, or (iii) both. Key to disentangling these interpretations, is the deployment of a spatially "neutral" pre-cue condition which permits the assessment of both the orienting component (via a comparison of valid versus neutral trials) and the reorienting component (via a comparison of invalid versus neutral trials).

The current study therefore aimed to investigate developmental changes in the deployment of selective attention during visual search, from childhood to adulthood. Moreover, the current developmental design also aimed to yield insight into the mechanisms of attention in general (i.e., in mature adults as well as other age groups; see Enns et al., 1998; Matusz et al., 2015) by demonstrating how the interplay between voluntary and reflexive components of selective attention is affected by developmental changes in the control of selective attention. We presented 6-, and 10-year-old children and young adults (20-year-olds) with visual scenes. Each scene included several objects, but only one agentive element (an animal or a human) which was the target stimulus, located on the right or left side of the scene. The participants were asked to search for the target (i.e., a voluntary attention task). Each scene was preceded by a sound (i.e., a burst of white noise serving as an auditory[1] pre-cue; duration = 50 ms) that was presented 50 ms before the scene. The auditory pre-cue was equiprobably presented either on the target side (valid trials), on the opposite side (invalid trials), or on both sides (neutral trials). Consequently, the auditory pre-cue was entirely task-irrelevant, that is, it was not a useful strategy to voluntarily pay attention to the auditory cue in order to carry out the visual search task. This cuing procedure therefore enabled us to disentangle the specific contributions of the orienting (indexed by valid minus neutral trials) and reorienting (invalid minus neutral trials) components of reflexive crossmodal attentional mechanisms during the voluntary searching task.

Studies of visual search typically use very simple, repetitive and unrealistic displays. However, there is now a growing consensus among researchers (and increasing efforts are being made in this respect) that it is necessary to use more naturalistic stimuli and behavioral tasks in order to gain a valid picture of how spatial attention works in ecological situations (for reviews see Felsen & Dan, 2005; Peelen & Kastner, 2014). For this reason, the visual searches we asked participants to make were to complex and unrepeated cartoon-like scenes. The use of complex scenes also allowed us to characterize the impact of visual features included in each scene, thus to measure also reflexive orienting towards task-relevant or task-irrelevant visual elements. To do this, we computed a "saliency map" for each scene using a well-known algorithm (Itti, Koch & Niebur, 1998) based on local discontinuities in line orientation, intensity contrast, and color opponency (see Fig. 1B). Saliency maps are topographic representations of dominant locations in a visual scene that are more likely to be attended to and processed during scene exploration (see, for a review, Itti & Koch, 2001). Each dominant (or salient) location in the scene is represented by a saliency "peak", that is, the highest saliency value within that location. The higher the value of the location's saliency peak, the more likely it is that that location will be attended to and processed during scene exploration.

By means of the above-described saliency maps, we computed two saliency-related indexes for each scene. The first one was based on the computation of the "number of salient locations" (NSL) within each scene. We used this index to characterized the dispersion of attentional resources over locations that are not necessarily involved with the detection of the to-be-searched target. If searching performance decreased (i.e., if we observed longer reaction times) as a function of an increased number of salient locations, this would mean that participants waste spatial attentional resources on salient but task-unrelated locations. In this sense, this index is a measure of an inefficient search strategy. The second saliency-related index was based on the level of perceptual saliency related to the current target stimulus, namely the "target saliency" (TgS) index, which corresponded to the saliency peak falling inside the to-be-searched target in the scene. If searching performance increased (i.e., if we observed shorter reaction times) as a function of increased target saliency, this would mean that participants could successfully use task-relevant (i.e., target-related) sensory information to accomplish the searching task.

If age-related level of attentional control plays a critical role on the interplay between voluntary and reflexive mechanisms of selective attention, we would expect larger reflexive orienting and reorienting effects in children as compared to young adults, resulting from a lesser capacity to remain focussed on task-relevant target features while filtering out task-irrelevant information (Enns et al., 1998; Rueda et al., 2004; Waszak et al., 2010). Considering visual saliency, we expected that the different age groups would provide opposite effects in relation to the two saliency indexes. Specifically, we expected that the number of salient locations would predict visual search performance in children but not in adults, indicating a "dispersion" of attentional resources over the different salient locations irrespective of their task-relevancy. By contrast, we expected that target saliency would predict visual search performance in adults but not in children, demonstrating a greater capability to select task-relevant information based on low-level perceptual saliency.

2. Methods and materials

2.1. Participants

One hundred and nine healthy participants volunteered for and took part the study. The sample included: forty children attending the first

[1] The main reason for using the auditory modality for spatial cuing was to investigate simultaneously both auditory and visual distraction, implemented respectively in terms of "invalid" audio-spatial cues and visual distractors (i.e., non-target objects) in the scene. Future studies might verify whether visuospatial cues – presented on the periphery of the scene or cuing within-scene locations – have a different impact on searching performance on complex visual scenes.