Acta Psychologica 152 (2014) 75–83

Contents lists available at ScienceDirect

Acta Psychologica

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

Attentional cueing by cross-modal congruency produces both facilitation and inhibition on short-term visual recognition☆

Elena Makovac [a,b,*], Sze Chai Kwok [a], Walter Gerbino [b]

[a] Neuroimaging Laboratory, Santa Lucia Foundation, Rome, Italy
[b] Psychology Unit, Department of Life Sciences, University of Trieste, Italy

ARTICLE INFO

Article history:
Received 4 November 2013
Received in revised form 14 July 2014
Accepted 16 July 2014
Available online 13 August 2014

PsycINFO classification:
2346
2343
2320

Keywords:
Attention
Cross-modal binding
Inhibition of return
Inverse effectiveness rule
Multisensory integration
Visual working memory

ABSTRACT

The attentional modulation of performance in a memory task, comparable to the one obtained in a perceptual task, is at the focus of contemporary research. We hypothesized that a biphasic effect (namely, facilitation followed by inhibition) can be obtained in visual working memory when attention is cued towards one item of the memorandum and participants must recognize a delayed probe as being identical to any item of the memorandum. In every trial, a delayed spiky/curvy probe appeared centrally, to be matched with the same-category shape maintained in visual working memory which could be either physically identical (*positive* trials) or only categorically similar (*negative* trials). To orient the participant's attention towards a selected portion of a two-item memorandum, a (*tzk*/*wow*) sound was played simultaneously with two lateral visual shapes (one spiky and one curved). Our results indicate that an exogenous attentional shift during perception of the memorandum, induced by a congruent audio–visual pairing, first facilitates and then inhibits the recognition of a cued item (but not of a non-cued item) stored in visual working memory. A coherent pattern of individual differences emerged, indicating that the amount of early facilitation in congruent-sound trials was negatively correlated with recognition sensitivity in no-sound trials (suggesting that the inverse effectiveness rule may also apply to memory) and positively correlated with later inhibition, as well as with the self-reported susceptibility to memory failures.

1. Introduction

Competition for representational resources makes attention a crucial factor in both perceptual and reflective domains, supporting comparisons and analogies between external and internal attention that are particularly relevant for research on working memory, which lies at their intersection (Chun, Golomb, & Turk-Browne, 2011). Attention affects encoding (Eger, Henson, Driver, & Dolan, 2004) as well as memory retrieval (Guerin, Robbins, Gilmore, & Schacter, 2012; Wolfe, Reinecke, & Brawn, 2006). Selection in the domains of visual perception and visual memory appears to be supported by similar neural mechanisms (Astle, Scerif, Kuo, & Nobre, 2009; Kuo, Rao, Lepsien, & Nobre, 2009; Nobre et al., 2004). Spatial memory is distorted by a task-irrelevant exogenous cue (Van der Stigchel, Merten, Meeter, & Theeuwes, 2007), and object memory deteriorates when attention shifts away from object location during retention (Williams, Pouget, Boucher, & Woodman, 2013).[1]

When observers are required to detect a target following the presentation of an exogenous cue that attracts spatial attention towards its location, the manipulation of *cue–target onset asynchrony* (CTOA) reveals a characteristic biphasic effect of cueing (Posner & Cohen, 1984). Detection is facilitated at short CTOAs (typically, less than 300 ms), while it deteriorates at longer CTOAs, generating an effect labeled as *inhibition of return* (IOR).

Classic IOR – that is, the detection loss for visual targets displayed in a previously attended location after a critical CTOA – can be explained by the *reorienting hypothesis*, which states that attention is automatically attracted towards the location of the peripheral cue (i.e., a lateral flash), but is subsequently disengaged from that particular location, because of a compensatory mechanism that inhibits the return of attention to previously attended locations, to maximize efficiency of visual search in a normally complex environment (Danziger, Kingstone, & Snyder, 1998; Klein, 2000). For an alternative, motor-based, view of IOR, as well as for its possible occurrence in a perceptual discrimination task, see Taylor and Donnelly (2002).

Attention can be attracted towards locations, objects, and features (Carrasco, 2011). Accordingly, IOR is not unique to spatial attention, having been found also when attention is object-based (i.e., oriented to targets that belong to a previously attended object; List & Robertson, 2007) and feature-based (i.e., oriented to targets that possess a previously attended feature; Busse, Katzner, & Treue, 2006). Outside the visual

☆ Authors thank David Pearson for comments on an earlier version of the manuscript. Supported by Miur-Cofin 2008 (PRIN200879EB93) and EcoAutobus-Industria (MS01_00006) 2015 grants.
* Corresponding author at: IRCCS Santa Lucia Foundation, Neuroimaging Laboratory, Via Ardeatina 306, Rome, Italy. Tel.: +39 3382305385.
[1] For a recent review on how selective attention can operate on perception, memory, and imagery in equivalent ways see Gosling and Astle (2013).

http://dx.doi.org/10.1016/j.actpsy.2014.07.008
0001-6918/

domain, IOR also occurs in auditory and audio–visual domains, provided that a second cue redirects attention back to the center (Reuter-Lorenz & Rosenquist, 1996).

Based on the idea that external and internal types of attention may share similar mechanisms, attentional cueing effects have been recently investigated in memory. Besides the well-known facilitatory effect of cueing on encoding (Uncapher, Hutchinson, & Wagner, 2011), attention can improve the maintenance of objects in memory and increase the probability of their recall (Murray, Nobre, Clark, Cravo, & Stokes, 2013). Johnson et al. (2013) manipulated participants' internal attention by presenting two items followed by a cue that required them to selectively think back to (i.e., *refresh)* only one item. Slower responses for refreshed than unrefreshed items revealed an IOR-like effect within working memory. In this case, internal attention was directed towards the semantic meaning of a word, which was independent of the memory of its spatial location. This again suggests that the memory-based IOR-like effect is not exclusively linked to spatial attention, but can arise from semantic cueing, as described in the perceptual domain (Fuentes, Vivas, & Humphreys, 1999).

Cross-modal congruency refers to the correspondence between inputs to different modalities that can make sounds and visual shapes perceptually similar (Köhler, 1929). Following previous demonstrations that cross-modal congruency can improve recognition (Murray et al., 2004), speed up cross-modal event detection (Makovac & Gerbino, 2010), and modulate attention (Chiou & Rich, 2012), here we asked whether the exogenous cueing of a perceptual event (i.e., the automatic orienting of external attention towards one of two visual shapes) can influence recognition. We expected that cross-modal congruency, by promoting multisensory integration and affecting the deployment of attention, would generate a biphasic effect of cueing on recognition performance. In particular, we explored the possibility that *visual working memory* (VWM) displays a *biphasic effect of attentional cueing* (facilitation followed by inhibition). Participants in our experiment were shown one spiky and one curvy shape on the left/right of the fixation point and maintained them in VWM until a central probe prompted for an *old*/*new* response based on physical identity. A *new* response (negative trials) was required when the match between the probe and the item of the memorandum with the same contour type was categorically similar (but not identical), whereas an *old* response was required when an identical probe was presented.

Our paradigm included multisensory and unisensory conditions. In multisensory conditions the memorandum was presented together with a simultaneous sound whose auditory features were congruent with the features of one of the two visual shapes, giving rise to cued trials. Cued trials were either valid (if sound and probe were congruent) or invalid (if sound and probe were incongruent). The unisensory condition included uncued trials (neutral; no sound was presented). Importantly, the sound in the multisensory conditions was task-irrelevant, thus making cued and uncued (neutral) trials formally equivalent in terms of task demand. Expected effects of cueing (facilitation followed by inhibition) should be dependent on *automatic cross-modal binding*.

We assumed that: (*a*) cross-modal congruency promotes the automatic binding of the sound with only one visual shape; (*b*) attention is exogenously oriented towards the cued shape; (*c*) multisensory integration enhances encoding (Lehmann & Murray, 2005; Nyberg, Habib, McIntosh, & Tulving, 2000; Wheeler, Petersen, & Buckner, 2000) by driving exogenous attention towards the multisensory event comprising congruent sound/cued shape combination (Spence, McDonald, & Driver, 2004; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010); and (*d*) the central probe produces a disengagement of attention from the peripheral cued shape, and gives rise to IOR (Posner & Cohen, 1984).

We expected the following two results:

1) *Facilitation*; when the probe is displayed immediately after the memorandum, recognition in valid trials should be facilitated, because the exogenously cued shape should benefit from enhanced encoding, at the expense of poorer encoding of the uncued shape;
2) *Inhibition*; at a longer probe delay (around 1 s), recognition in valid trials is inhibited, as revealed by a reduction in recognition performance for the exogenously cued shape (Lupiáñez, Milán, Tornay, Madrid, & Tudela, 1997; Massen & Stegt, 2007).

We also explored the correlation between individual differences in performance and the self-reported frequency of cognitive mistakes, measured by the Memory and Distractibility subscales of the Cognitive Failures Questionnaire (CFQ; Broadbent, Cooper, FitzGerald, & Parkes, 1982; Wallace, Kass, & Stanny, 2002). We hypothesized that the amounts of the expected facilitation and inhibition effects (which should reveal how attention operates in VWM) could be positively correlated with the awareness of the propensity to attention and memory errors in everyday life, as measured by the relevant CFQ items. The reliability and validity of CFQ in quantifying the propensity for making mistakes have been extensively studied (Forster & Lavie, 2007; Kanai, Dong, Bahrami, & Rees, 2011; Martin & Jones, 1983; Tipper & Baylis, 1987). However, CFQ scores might also reflect metacognitive worries (Mecacci & Righi, 2006) and the tendency to pessimistic self-evaluations (van Doorn, Lang, & Weijters, 2010).

2. Method

2.1. Participants

Twenty right-handed undergraduates (14 females, mean age = 24 years, age range 19–29, SD = 3.5) participated in the experiment. All participants had normal hearing and normal/corrected-to-normal visual acuity. All gave their prior informed consent, were tested individually, and received course credit.

2.2. Stimuli, apparatus, and procedure

As regards visual stimuli, inspired by Köhler (1929, Figs. 18 and 19) we generated 40 unfamiliar shapes — 20 with spiky contours containing abrupt and frequent (median = 28 in the 14–44 range) discontinuities, and 20 with curvy contours containing smooth and infrequent (median = 8 in the 4–12 range) changes of curvature polarity (i.e., from convexity to concavity and vice versa). They were drawn manually, adding variable portions (spiky vs. curvy) to a central disk subtending about 3° and reaching a maximum angular extent of about 9.0° at presentation (Fig. 1A) (see Supplementary material for the whole set of visual shapes). We were aware that spiky and curvy shapes differed in their complexity, if this is defined as a function of the number of contour changes (Baylis & Driver, 2001). The shapes appeared light gray (20.7 cd/m^2) on a black screen (5 cd/m^2). Nine practice pairs and 288 experimental pairs of visual shapes, to be used as memoranda, were extracted from the 400 spiky–curvy pairs, balancing for position.

As regards acoustic stimuli, we used Audacity® 2.0 to generate two 200-ms sounds: *tzk*, a spiky sound with abrupt and frequent changes of intensity, and *wow*, a soft sound with smooth and infrequent changes of intensity.[2]

Participants were seated 57 cm away from a dimly illuminated screen and instructed to pay attention to visual memoranda, treat sounds as task irrelevant, and prioritize accuracy over speed when matching the memorandum to the probe for physical identity. After a 10-min dark adaptation period and a 9-trial practice block, four 72-trial blocks were presented, lasting about 12 min each, with 3-min breaks between blocks. Every trial started with a 1000-ms green fixation cross displayed at the center of the screen, followed by the onset of a 150-ms memorandum made of two shapes (with their centers at ±7.5° of

[2] The audio files "tzk_C.wav" and "wow_C.wav" are available as Supplementary material.