Vision at a glance: The role of attention in processing object-to-object categorical relations

Nurit Gronau¹

Published online: 6 January 2020
© The Psychonomic Society, Inc. 2020

Abstract

When viewing a scene at a glance, the visual and categorical relations between objects in the scene are extracted rapidly. In the present study, the involvement of spatial attention in the processing of such relations was investigated. Participants performed a category detection task (e.g., “is there an animal”) on briefly flashed object pairs. In one condition, visual attention spanned both stimuli, and in another, attention was focused on a single object while its counterpart object served as a task-irrelevant distractor. The results showed that when participants attended to both objects, a categorical relation effect was obtained (Exp. 1). Namely, latencies were shorter to objects from the same category than to those from different superordinate categories (e.g., clothes, vehicles), even if categories were not prioritized by the task demands. Focusing attention on only one of two stimuli, however, largely eliminated this effect (Exp. 2). Some relational processing was seen when categories were narrowed to the basic level and were highly distinct from each other (Exp. 3), implying that categorical relational processing necessitates attention, unless the unattended input is highly predictable. Critically, when a prioritized (to-be-detected) object category, positioned in a distractor’s location, differed from an attended object, a robust distraction effect was consistently observed, regardless of category homogeneity and/or of response conflict factors (Exp. 4). This finding suggests that object relations that involve stimuli that are highly relevant to the task settings may survive attentional deprivation at the distractor location. The involvement of spatial attention in object-to-object categorical processing is most critical in situations that include wide categories that are irrelevant to one’s current goals.

Keywords Visual attention · Object–object associations · Focused attention · Categorical processing · Task relevance · Contingent attention capture · Perceptual categorization and identification · Object recognition

What do we grasp during a brief glimpse? A “snapshot” of the world typically involves colorful and cluttered sights that require rapid segmentation of objects from their background and from each other. Limited by resource capacity constraints, our cognitive system fails to register all objects and the abundance of their visual detail at a glance. One factor that may moderate visual complexity and streamline object recognition, however, is the network of semantic and visual associations existent between an individual object and its larger contextual surrounding. A truck, for instance, normally appears on a road, near some cars, street signs, and/or traffic lights, within a broader setting of a street or a highway. On the basis of global scene properties, rapid recognition of the scene’s gist may generate a limited set of

predictions with regard to expected object identities, thus facilitating their recognition (Antes, Penland, & Metzger, 1981; Bar, 2004; Biederman, 1972, 1981; Boyce, Pollatsek, & Rayner, 1989; Davenport & Potter, 2004; Friedman, 1979; Oliva & Torralba, 2007; Palmer, 1975; but see different approaches to scene–object facilitation in Demiral, Malcolm, & Henderson, 2012; Ganis & Kutas, 2003; Hollingworth, 1998). Similarly, recognizing a central object within a visual setting, such as a truck, may activate or prime other nearby associated objects, thereby reducing the cognitive resources required for their processing and facilitating the formation of a coherent visual percept. Indeed, when stripping off background information and focusing on local object-to-object relations, semantic, categorical, and visual associations among individual items are efficiently extracted when stimuli are only glanced for a very brief duration (e.g., Auckland, Cave, & Donnelly, 2007; Biederman, Mezzanotte, & Rabinowitz, 1982; Davenport, 2007; Dobel, Gumnior, Bölte, & Zwitserlood, 2007; Green & Hummel, 2006; Gronau & Shachar, 2014; Henderson, Pollatsek, & Rayner, 1987;

✉ Nurit Gronau
nuritgro@openu.ac.il

¹ Department of Psychology & Cognitive Science Studies, Open University of Israel, Raanana, Israel

Oppermann, Hassler, Jescheniak, & Gruber, 2012; Roberts & Humphreys, 2011).

To what extent is visual attention involved in the extraction of object-to-object relations, and more generally, in the formation of a coherent visual percept? Traditional attention models have posited that at any moment only one or perhaps a very small number of objects can be attended and fully processed (e.g., Kahneman, Treisman, & Gibbs, 1992; Treisman & Gelade, 1980; Wolfe & Cave, 1999). Supposedly, the attended stimuli can be linked together (or dissociated from each other), depending on items' visual and semantic relations. Objects appearing outside the main focus of attention, however, are largely filtered out and/or are processed only to a coarse level (e.g., Evans & Treisman, 2005). Salient changes to unattended objects may be overlooked and the construction of a global scene representation may be incomplete (e.g., Mack & Rock, 1998; Rensink, O'Regan, & Clark, 1997; Simons & Levin, 1997). According to classical approaches, thus, analyzing the relations between unattended stimuli, or between attended and unattended objects, likely requires an explicit shift of visual attention to the unattended regions in the scene.

One exception to this filtering, or attenuated processing of stimuli appearing outside the main focus of spatial attention, is the case of items that possess features that are highly relevant to the current task requirements. Ample evidence has demonstrated that when engaging in a search, or in a detection of a certain target, other distractor stimuli sharing visual and/or conceptual features with this target may be processed in parallel (e.g., Duncan & Humphreys, 1989; Peelen, Fei-Fei, & Kastner, 2009; Seidl-Rathkopf, Turk-Browne, & Kastner, 2015; Wolfe, Cave, & Franzel, 1989) or may elicit an involuntary shift of spatial attention to their location (Bacon & Egeth, 1994; Folk & Remington, 1998; Folk, Remington, & Johnston, 1992; Wyble, Folk, & Potter, 2013). The beneficial status of stimuli positioned in unattended locations but that carry features (or feature dimensions) that comply with one's task settings is the hallmark of theories emphasizing feature-based attention (Hillyard & Münte, 1984; Maunsell & Treue, 2006), feature dimension weighting (e.g., Memelink & Hommel, 2013; Müller, Heller, & Ziegler, 1995; Müller, Reimann, & Krummenacher, 2003), and/or contingent attention capture (Folk & Remington, 1998; Folk et al., 1992).

A more intriguing case, however, is the one in which stimuli are not actively searched for; hence, they do not necessarily comply with one's current goals. Recent research has shown that under unattended, or minimally attended conditions, processing the semantic relations between pairs of *non-prioritized* (e.g., nonsearched) objects is in fact reduced to minimum. In one study, for instance, participants were required to detect a nonsense shape among pairs of stimuli containing either nonsense or real-world objects. When participants were cued to respond to one of two semantically associated (real-world) objects (e.g., a cued pizza cutter presented nearby an uncued pizza), no

relational effects were observed among the stimulus pairs. This was in stark contrast to a situation in which spatial attention was spread across both objects, and their semantic relations, though irrelevant to task requirements, affected performance significantly. Focusing attention on one object only, thus, eliminated the relational effects seen when both objects were fully attended (Gronau & Shachar, 2014). Similarly, when semantically associated object images served as irrelevant backgrounds to target letters, participants tended to dwell on pairs of related, rather than on unrelated images. This semantic bias, however, was contingent on the objects being presented for long exposure durations and on attention being explicitly directed to them (e.g., Malcolm, Rattinger, & Shomstein, 2016). These findings are in accordance with classical attention models arguing for the necessity of visual attention to the identification of complex objects and their relations (e.g., Treisman & Gelade, 1980).

Importantly, objects appearing within real-world visual settings often (albeit not always) belong to the same overall category, resulting not only in semantic proximity (i.e., a common global context) but also in a relative perceptual similarity among the objects. A sofa in a living room is typically seen near a couch or a table, and a truck on the road is viewed near some other cars or vehicles. Since an attended truck shares more visual features with a car than with a traffic light or a street sign, this relative similarity among vehicles may aid processing of the car when appearing in an unattended region, even when the latter is irrelevant to one's current goals (e.g., when searching for a human figure). That is, unattended objects that do *not* belong one's explicit task setting may nevertheless be processed to a significant degree due to (partial) activation of their features by attended items. The present study focused on such everyday situations, in which an attended object is surrounded by items that belong to the same superordinate category, but are not necessarily exemplar "copies" of that very object.

Note that visual similarity has been shown to serve as an important factor in many paradigms examining stimulus relations and their effects on item detection and identification. An extreme case of item similarity is the appearance of identical stimuli, characterized by full featural overlap. In the redundancy gain (RG) paradigm, for instance, the presentation of two identical (or highly similar) stimuli yields faster item detection than the presentation of a single stimulus (e.g., Holmgren, Juola, & Atkinson, 1974; Mordkoff & Yantis, 1991; Van der Heijden, La Heij, & Boer, 1983). These stimuli, however, are defined *a priori* as an inherent part of one's task setting (i.e., to-be-detected targets) and are presented within the main focus of attention (often, to the left and right of fixation). Any stimulus similarity effects observed with the RG, thus, may be associated with the fact that stimuli are highly prioritized, and are effectively attended. Likewise, the effects of similarity among targets and distractors (as well as among the distractors themselves) have been extensively investigated within the visual search framework (e.g., Duncan & Humphreys, 1989; Jacobs, 1986; Roper,

Cosman, & Vecera, 2013). Here too, stimuli are typically attended, or partially attended, as spatial attention freely shifts within the display, allowing individual items to be brought to its main focus. In addition, guidance of attention is mediated by featural and perhaps conceptual properties of the searched target. Target–distractor similarity effects, thus, are influenced to a large extent by goal-directed attentional settings. Other paradigms, such as the well-known flanker task, have also examined the effects of target–distractor similarity on the flanker distraction effect (e.g., B. A. Eriksen & Eriksen, 1974; C. W. Eriksen & Eriksen, 1979; Fournier, Scheffers, Coles, Adamson, & Abad, 1997). In contrast to visual search paradigms, spatial attention in the flanker task is focused in advance on a preknown (e.g., central) location, while distractors are positioned “outside” the main focus of attention. Flankers resembling the target, however, typically belong to a limited set of to-be-detected items (either congruent, or incongruent with the target), thus, by definition these distractors are prioritized by task demands. Once again, any influence of the peripheral flankers on the central target may be related to their prioritization by online task settings.

It appears, then, that although several paradigms have examined the effects of visual similarity on item processing, such effects remain largely uninvestigated within the context of *unattended* stimuli that are *not* prioritized by the task requirements. As was mentioned earlier, the distinction between prioritized and nonprioritized stimuli is highly important within the framework of studies investigating “unattended” processing, as stimuli positioned outside the main focus of spatial attention may in fact be activated and selected for high-level processing via feature-based or other goal-directed attentional mechanisms (Folk et al., 1992; Maunsell, & Treue, 2006; Memelink & Hommel, 2013; Müller et al., 1995; Wyble et al., 2013; see also Gronau, Cohen, & Ben-Shachar, 2003, 2009; Gronau & Izoutcheev, 2017). It is under the more stringent conditions of *irrelevance* to the current settings that visual and categorical similarity effects are of interest and should be examined.

In addition, the perceptual variance among visually similar (or dissimilar) stimuli in previous research has typically been very limited, as most studies have used a small set of arbitrary (i.e., meaningless) and/or symbolic stimuli. Stimulus variability in these studies, thus, does not reflect the visual richness and the abundance of object types characterizing items in real-world scenes.

The present study, therefore, focused on the involvement of spatial attention in processing visual and semantic relations among real-world objects sharing the same superordinate category. Specifically, I examined whether visual/categorical effects (e.g., faster latencies to nonprioritized objects belonging to the same relative to different superordinate categories) existed when attention was focused on one object while a categorically related object was positioned at the outskirts of focal attention. Since categorical processing of individual items is known to take place rather instantaneously (e.g., Clarke & Tyler, 2015; Grill-Spector & Kanwisher, 2005;

Thorpe, Fize, & Marlot, 1996; see the review in Fabre-Thorpe, 2011), I wished to explore whether categorical relations among stimuli may affect behavior during a very brief glance, despite the underprivileged status of the unattended object. Namely, I investigated the question of whether, in the absence of a clear feature-based (and/or conceptual-based) attentional setting, activation of an unattended object by a categorically related item was possible, leading to categorical effects similar to the ones observed when both objects are attended (see below).

To examine this question, participants were briefly presented with object pairs that were taken from the same or from different superordinate object categories (e.g., two vehicles vs. a vehicle and an animal, respectively), and the responses to these relations were measured in an implicit manner. The participants were instructed to detect a third category, composed of either meaningless (Exp. 1A) or real-world (Exp. 1B) objects, that was defined as the “target” category, while the other two categories were considered (unbeknownst to participants) as “nontarget” (nonprioritized) categories. In the basic experimental version, participants’ task was to determine rapidly, on each trial, whether the target category was present or not in the display. It was hypothesized that regardless of the nature of images (i.e., whether they contained a target category or not), same-category pairs would be processed faster and more accurately than different-category pairs. This hypothesis was based on two main lines of research. The first involved perceptual similarity effects on processing, detection, and identification of arbitrary and/or symbolic stimuli (e.g., during visual search, Duncan & Humphreys, 1989; Jacobs, 1986; Roper et al., 2013). The second line of research involved images of real-world objects, demonstrating priming effects among pairs of pictorial stimuli belonging to the same superordinate category (e.g., Bajo, 1988; Carr, McCauley, Sperber, & Parmelee, 1982; Henderson et al., 1987; Kroll & Potter, 1984; Sperber, McCauley, Ragain, & Weil, 1979), as well as object–object facilitative contextual and categorical effects (e.g., Auckland et al., 2007; Davenport, 2007), even when the stimuli were irrelevant to the current task requirements (e.g., Gronau & Shachar, 2014; Malcolm et al., 2016). Note that in most of the cases mentioned above, the stimuli were attended and/or were actively brought to the main focus of visual attention (via search). In accordance with these studies, the images in both locations in the basic experimental version (i.e., Exp. 1) were considered attended, since the target category could appear in either one of the locations and spatial attention was likely to span both stimuli. Once a “categorical relation effect” was obtained among the nontarget stimuli under these attended conditions (i.e., streamlined responses to same- than to different-category trials), one could proceed to examining the necessity of visual attention to processing these categorical relations. Namely, the very same effect was examined under conditions in which one of two images was

irrelevant to task requirements and was positioned outside the main focus of visual attention. Thus, in the “unattended” experimental version (Exp. 2), participants focused on, and responded to, one of two images, while its counterpart image functioned as an irrelevant distractor. This paradigm resembled a flanker task, yet in contrast to typical flanker paradigms, the nontarget stimuli examined were *not* prioritized by task demands (i.e., were not part of the target set), and their categorical relations were independent of the response congruency/redundancy (or response inhibition) factors. To the extent that a meaningful categorical relation effect was observed under these conditions, one could infer that the processing of stimuli’s categorical relations did not critically rely on focal attention. That is, object-to-object visual and/or conceptual associations were processed despite the underprivileged status of the unattended distractor. If, however, no categorical relation effect was observed when one of two stimuli was deprived of focal attention, one could reasonably conclude that the extraction of visual/conceptual relations was not an attention-free process. Using traditional statistical analyses as well as a Bayesian approach, firm evidence could be provided for a null result in such a case.

As a preview to the following studies, Experiments 3A and 3B were follow-ups to Experiment 2, in which the visual similarity within and between categories was further manipulated under “unattended” conditions. In Experiment 4, the impact of task-related settings on object-to-object visual/categorical processing was examined more closely.

Experiment 1: Both object locations are relevant for task demands

Method

Participants Two experiments were conducted. In Experiment 1A participants were instructed to detect a meaningless shape (following Gronau & Shachar, 2014), whereas in Experiment 1B they were to detect a real-world target category (thus minimizing perceptual differences between target and nontarget categories). Twenty-one¹ undergraduate students with normal or corrected-to-normal sight participated

in each of the experiments for course credit (Exp. 1A: 15 females, six males; ages 19–43 years; Exp. 1B: 10 females, 11 males; ages 19–35 years). The experiments (as well as all following studies) have been approved by the Israel’s Open University Ethical Review Board committee. All participants signed a consent form indicating that participation was voluntary and that they could withdraw from the experiment at any time without penalty.

Apparatus Stimulus presentation and response collection were controlled by a PC computer, using an E-Prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2002). Stimuli were presented on a 15-in. CRT monitor with an 85-Hz refresh rate.

Experimental design Each trial contained a pair of stimuli presented simultaneously above and below fixation. As was mentioned above, in Experiment 1A participants were instructed to detect a meaningless shape among pairs of items comprising either meaningless target shapes, real-world (nontarget) objects, or both (see examples of meaningless stimuli in Fig. 1). In Experiment 1B the target items belonged to a real-world superordinate category (e.g., clothes), differing from the other two nontarget categories (animals, transportation). The target category was counterbalanced across participants. The experimental design was identical in both experiments: Among trials containing nontarget stimuli (240 trials), half of the pairs contained same-category objects, whereas the other half contained different-category objects (see the examples in Fig. 2a and b, respectively). In addition, 210 pairs of stimuli containing a target item were similarly divided into a same-category condition (two target stimuli, 90 trials), and a different-category condition, comprising a target stimulus paired with a nontarget object (120 trials; see Fig. 2c and d, respectively). Ninety additional trials containing a single item were used. These consisted of 30 target and 60 nontarget stimuli, appearing either at an upper (50%) or a lower location. The single-item trials were particularly important in Experiment 2 (for assessing cue-validity effects, see below), and were inserted here in order to allow similar presentation conditions across experiments.

Stimuli Object stimuli were presented on a gray background square subtending about 22 cm × 22 cm (corresponding to a visual angle of approximately 20° by 20° from a viewing distance of 60 cm). Objects’ size varied, ranging from 3.0 to 6.5 cm (visual angle of approximately 2.8°–6°), in both height and length dimensions, and centered approximately 3.5° above/below fixation. The images were gathered from various sources, including commercially available CDs and the Internet. The meaningless shapes serving as targets in Experiment 1A were created via Photoshop software. Note that within the real-world object categories, although the main interest was in visual and categorical relations that allow, by

¹ According to power computations, to achieve a power of .9 under the assumption of a categorical relation effect of $d = 0.8$ ($\alpha = .05$), a sample of 19 participants was required. Since there were three target categories in Experiment 1B, and to allow for full target counterbalancing across participants, a sample of 21 participants was chosen for both studies of Experiment 1, as well as for the following experiments (aside from Exp. 4, in which a slightly larger sample was required, due to a change in counterbalancing requirements). In addition to these a priori power computations, a post-hoc computation of the achieved power was conducted for the categorical relation effect among the nontarget trials of Experiment 1 (see below). The power computation yielded an actual power of .98 in Experiment 1A, and a power of .93 in Experiment 1B. All power computations were conducted using the G*Power 3.1.3 software (Faul, Erdfelder, Lang, & Buchner, 2007).

Fig. 1 Examples of the meaningless target shapes used in Experiment 1A

definition, featural overlap among exemplars belonging to the same category, obvious repetitiveness of specific exemplar images within each category were avoided as much as possible. That is, I wished to prevent participants from actively tuning themselves to highly specific visual features and/or object configurations. This was particularly important in Experiment 2, in which one of two object images served as an irrelevant distractor. I therefore used rather wide superordinate categories that were composed of a range of basic-level categories. Such superordinate categories were previously shown to be detected “ultra” rapidly, even when objects were embedded within a natural, cluttered background (e.g., Thorpe et al., 1996; see the review in Fabre-Thorpe, 2011). The superordinate category of animals, therefore, included images of mammals, birds, fish, reptiles, amphibians and insects; the category of transportation/vehicles included images such as cars, trucks, carriages, bikes, buses, boats and planes;

and the category of clothes (used in Exp. 1B) included images such as shirts, trousers, jackets, dresses, skirts, coats, and various accessories. For masking purposes, an achromatic texture image depicting a “pseudo-noise” pattern was used, taken from a database of meaningless achromatic texture images (see the Procedure section and Fig. 3). The root mean square contrast (defined as the standard deviation of the pixel intensities) of the mask image was 0.12.

Procedure Each trial lasted approximately 2 s (see Fig. 3). The trial began with a fixation cross appearing at the screen center for 547 ms, followed by a pair of object stimuli presented for 24 ms. Subsequently, a grayscale pseudo-noise pattern mask image appeared for 129 ms, followed by a 1,082-ms blank lasting until the end of the trial (the same mask was used throughout all trials in the experiment; see Fig. 3). Participants were instructed to maintain fixation in the course of all trials.

Fig. 2 Examples of the paired images within the different conditions in Experiment 1A. Note that the target category in the present example is *meaningless shapes* (replaced by a real-world object category—e.g., clothes—that was counterbalanced across participants in Exp. 1B). *Left panel:* Nontarget (i.e., target-absent) trials. (a) Same-category condition: Both objects are means of transportation, or both are animals. (b) Different-category condition: a vehicle and an animal. *Right panel:*

Target (i.e., to-be-detected category) trials. (c) Same-category condition: Both images are meaningless shapes (clothes). (d) Different-category condition: One of the two images depicts a meaningless shape (or a clothing article, in Exp. 1B), while its counterpart image depicts an animal or a vehicle. Within the different-category target trials, half of the trials contained a target in an upper location, and the other half a target in a lower location. All images were presented in color

Fig. 3 An example of a trial sequence in Experiment 1

Prior to the beginning of the experiment the participants were introduced with all categories and were notified about the identity of the target category. Their task was to determine whether the stimulus display contained a target stimulus or not by using a two-alternative forced choice key press (using the keys “J”/“K,” respectively). Participants were required to respond as fast as possible, while maintaining accurate performance (there was no feedback for correct or incorrect performance during the experiment). Importantly, within the trials containing a target stimulus, the target could appear in an upper or a lower location with equal probability, thus, participants had to be attentive to both stimulus locations. As mentioned earlier, these target trials were of minor interest, particularly in Experiment 1A in which nonsense stimuli were used. Furthermore, even in Experiment 1B, in which real-world objects served as to-be-detected targets (in an attempt to reduce perceptual differences between target and nontarget categories), any categorical relation effect among the target trials could potentially be confounded with an RG effect (e.g., Holmgren et al., 1974; Mordkoff & Yantis, 1991; Van der Heijden et al., 1983). This was due to the fact that the same-category condition contained two targets whereas the different-category condition contained only a single target item. In addition, within the “unattended” experimental version (i.e., Exp. 2, see below), this state of affairs would potentially cause a flanker-like situation, in which response activation patterns in the attended and the unattended locations would either be congruent (i.e., same category) or incongruent (different category) with each other. Any categorical relation effect under these presentation conditions, thus, could be accounted for by response congruency (redundancy) factors, rather than by pure visual and/or categorical associative relations (see attempt to overcome this problem in Exp. 4). The statistical analysis of the categorical relation effect in Experiment 1, thus, was conducted within the Nontarget trials only, which were the focus of the present research.

Each image within the different target and nontarget conditions appeared only once in the course of the experiment.

The order of the trials was determined randomly. Prior to the beginning of the experiment, there was a practice session of 56 trials. The object images in the practice session were excluded from the experiment.

Results and discussion

Mean response times (RTs) and accuracy rates were computed for each condition. Trials on which participants made errors (Exp. 1A, 13%; Exp. 1B, 10%) or trials yielding extreme outlier responses (i.e., RTs that deviated from the participant’s mean RT by more than three standard deviations, approximately 1% in both experiments), were excluded from the RT analysis. Because the responses in the single-item condition were of minor importance (Exp. 1A: RTs: $M_{\text{Targets}} = 535$, $SE_{\text{Targets}} = 10$, $M_{\text{Nontargets}} = 559$, $SE_{\text{Nontargets}} = 15$; Accuracy: $M_{\text{Targets}} = 0.88$, $SE_{\text{Targets}} = 0.02$, $M_{\text{Nontargets}} = 0.88$, $SE_{\text{Nontargets}} = 0.02$; Exp. 1B: RTs: $M_{\text{Targets}} = 505$, $SE_{\text{Targets}} = 13$, $M_{\text{Nontargets}} = 526$, $SE_{\text{Nontargets}} = 12$; Accuracy: $M_{\text{Targets}} = 0.89$, $SE_{\text{Targets}} = 0.01$, $M_{\text{Nontargets}} = 0.91$, $SE_{\text{Nontargets}} = 0.02$), Table 1 presents only the responses to trials containing object pairs. As was mentioned earlier, I focus the analysis on the categorical relation effect (i.e., the difference between the different- and same-category conditions) among the nontarget trials, which contained only “nonprioritized” categories. Here, both conditions contained similar response activation patterns, since all object locations contained nontarget stimuli. Within these trials, thus, responses to the object classification task were orthogonal to the main question of interest and were independent of redundancy gain factors.

Note that within the nontarget trials, mean responses were computed across the two nontarget categories (e.g., transportation and animals in Exp. 1A). For the accuracy rate measure, all statistical analyses were performed on Arcsine transformed accuracy data (Hogg & Craig, 1995).

A two-tailed within-subjects planned comparison was conducted in order to assess the categorical relation effect among the nontarget trials in each of the experiments. This contrast was computed using a traditional hypothesis-testing approach as well as a Bayesian approach, allowing us to determine the extent to which the differences observed in participants’ responses supported H1 (i.e., a categorical relation effect) or H0 (no categorical relation effect).² The planned *t* test revealed a strong categorical relation effect for the nontarget trials with the RT measure [Exp.

² A Bayes factor was computed, representing the likelihood ratio of the data under the assumption of the presence of a categorical relation effect (referred to as the BF; see Dienes, 2011; Jeffreys, 1961). According to widely accepted benchmarks, BF values of 10 or greater indicate a strong effect, $3 < \text{BF} < 10$ suggests a moderate effect, $0.33 < \text{BF} < 3$ implies an inconclusive result, and $\text{BF} < 0.33$ indicates the absence of an effect (i.e., clear support for H0). Cohen’s *d* effect size, computed for all contrasts, is defined as the standardized mean difference between the two conditions. According to common rules of thumb, *d* values of 0.2, 0.5, and 0.8 stand for small, medium, and large effects, respectively (see J. Cohen, 1988). All statistical analyses were conducted using the JASP software (version 0.9.0.1).

Table 1 Mean reaction times (in milliseconds) and accuracy rates in the categorical classification task, among the pair object trials of Experiments 1A and 1B (standard errors are in parentheses)

	Nontarget Trials		Target Trials	
	<i>Same</i>	<i>Different</i>	<i>Same</i>	<i>Different</i>
Experiment 1A (Target: meaningless shape)				
Reaction time	541 (17)	555 (17)	519 (9)	550 (10)
Accuracy rates	.9 (.02)	.87 (.02)	.93 (.02)	.78 (.02)
Experiment 1B (Target: real-world category)				
Reaction time	513 (10)	525 (10)	479 (13)	517 (14)
Accuracy rates	.91 (.02)	.92 (.01)	.95 (.01)	.81 (.02)

1A: $t(20) = 4.18$, $p < .001$, Cohen's $d = 0.91$, $BF = 71$; Exp. 1B: $t(20) = 3.60$, $p < .002$, $d = 0.79$, $BF = 21.8$], yet no effect with the accuracy rate measure [Exp. 1A: $t(20) = 1.57$, $p > .1$, $d = 0.34$, $BF = 0.66$; Exp. 1B: $t(20) < 1$, $p > .1$, $d = 0.07$, $BF = 0.24$]. The results suggest that, at least with respect to response latencies, displays containing images from the same superordinate object category are processed more efficiently than displays containing stimuli from different categories. Obtaining a categorical relation effect among the nontarget trials indicates that participants were sensitive to object categorical relations and/or to their relative perceptual similarity, and they processed them regardless of immediate task requirements. These findings may resonate earlier findings obtained in the visual search literature, in which faster responses were observed in both target-present and target-absent trials when distractors were visually similar (Duncan & Humphreys, 1989). Note, however, that these visual search studies used symbolic stimuli (rather than real-world objects), that were essentially identical under the “similar” (i.e., homogeneous) presentation conditions.

As was mentioned in the introduction, obtaining a categorical relation effect among pairs of real-world stimuli forms a baseline for the following experiments, in which visual attention would be focused on one of two objects, while its counterpart object would serve as a distractor that appeared in an irrelevant visual location.

Experiment 2: One of two objects is an irrelevant distractor

Experiment 1 has established that displays containing objects from the same superordinate category are processed more efficiently (i.e., faster) than displays containing stimuli from different categories. Importantly, one can assume with a high level of confidence that images in Experiment 1 were attended, as target location in each trial was unknown and participants presumably spanned their attention across the two stimulus locations. The goal of Experiment 2 was to

assess whether turning one of the two stimuli into an irrelevant distractor influenced the processing of the object-to-object visual and categorical relations. To this end, attentional allocation was manipulated by presenting a peripheral spatial cue prior to stimuli appearance, which summoned participants' visual attention to one of the two stimuli within a pair. In addition, task instructions were altered, such that the object classification task was now performed on the cued image only (e.g., “does the *cued* location contain a meaningless shape/article of clothing, or not?”), instead of on both images within a pair (“does the *display* contain a meaningless shape/article of clothing, or not?”). The cued image, therefore, was spatially attended and task-relevant, whereas the uncued image effectively became an irrelevant distractor positioned at the outskirts of focal attention. All other experimental parameters were identical to those in the previous experiment. Note that by inserting the spatial cue prior to stimulus presentation (along with changing the task instructions), spatial uncertainty was eliminated, effectively turning the paradigm from a diffused-attention task, or a search-like task (with an alleged set size of 2), into a focused-attention one (with a single, preknown relevant location only). I asked whether a categorical relation effect would be obtained under such conditions among the target-absent (nontarget) trials. To the extent that visual and categorical object–object relations were implicitly processed, a categorical relation effect should be observed, similar to the one observed in the previous experiment. If, however, the allocation of spatial attention is a necessary condition for linking the two associated objects, no categorical effect should be observed.

Method

Participants As in Experiment 1, two experiments were conducted, one in which meaningless shapes (Exp. 2A), and another in which a real-world objects (Exp. 2B), served as targets. Twenty-one undergraduate students (13 females, eight males; ages 15–39 years) participated in Experiment 2A, and an additional 21 independent students (15 females, six males; ages 21–34 years) participated in Experiment 2B for course credit. All of the participants had normal or corrected-to-normal sight.

Experimental design and stimuli The stimuli in Experiments 2A and 2B were identical to those in Experiments 1A and 1B, respectively. The insertion of the cue prior to stimulus presentation, however, created a slightly different division into subconditions within some of the experimental conditions. First, whereas the cue always marked one of two objects within the pair-object trials (see Fig. 4), within the single-object condition it created a classical cue-validity situation in which its location either contained an object (i.e., valid-cue trials, 50%) or not (invalid-cue trials). This division into cued and

Fig. 4 Examples of the pair images within the different conditions in Experiment 2A (the target category in the present example is *meaningless shapes*). A black frame surrounding an images denotes the cue, which appeared prior to stimulus presentation. The cue disappeared once the images appeared, yet it is presented here for illustrative purposes.

uncued single objects afforded the opportunity to validate the spatial cuing manipulation. Note that the overall cue validity (i.e., its informativeness) was very high across all single and pair-object conditions, as the invalid trials formed only 8.3% of trials in the experiment. Second, within the trials containing two objects, the *different* condition of the target item trials was now subdivided into two subconditions: Half of the trials contained a cued target image and an uncued nontarget image (*target-cued* condition), whereas in the other half, an opposite state of affairs existed (*target-uncued* condition; see Fig. 4d, left and right panels, respectively). This division into two subconditions, and specifically the latter condition, in Experiment 2B, in which real-world target objects were used, enabled the examination of distraction effects caused by a to-be-detected target category appearing in an unattended location (see the detailed explanation below). All other parameters of the design were identical to those in Experiment 1.

Procedure The procedure was identical to that of Experiment 1, except for the insertion of a peripheral cue—a black, three-sided frame of 0.4° width, surrounding the upper or lower object location. The cue was presented for 47 ms, followed by the presentation of the object stimuli (24 ms) and the achromatic noise-pattern mask (129 ms). The short presentation durations of the cue and the image stimuli permitted the allocation of visual attention to the cued location while minimizing inadvertent attention shifts and eye movements. As mentioned earlier, participants responded to whether the *cued location* contained a target category or not. Due to the insertion of the cue, fixation duration at the beginning of the trial was shortened to 500 ms. All other parameters of the procedure were similar to Experiment 1.

It appeared with equal proportions in the upper and lower locations within each of the conditions. (a, b) Nontarget (i.e., target-absent) conditions. (c, d) Conditions containing a target stimulus. Among the *different-category* target trials, the target was either cued (left, “Target cued” condition) or uncued (right, “Target uncued” condition)

Results and discussion

Mean RTs and accuracy rates were computed for each condition. Trials on which participants made errors (Exp. 2A, 13%; Exp. 2B, 12%) or trials on which there were extreme outlier responses (approximately 1% in both experiments) were excluded from the RT analysis. A robust cue-validity effect observed among the single-object conditions validated the effectiveness of the spatial cuing manipulation (Exp. 2A: Invalid-cue condition: $M_{\text{Targets}} = 629$, $SE_{\text{Targets}} = 33$, $M_{\text{Nontargets}} = 640$, $SE_{\text{Nontargets}} = 27$; Valid-cue condition: $M_{\text{Targets}} = 527$, $SE_{\text{Targets}} = 15$, $M_{\text{Nontargets}} = 497$, $SE_{\text{Nontargets}} = 12$, $t(20) = 7.2$, $p < .001$, averaged across target and nontarget items; Exp. 2B: Invalid-cue condition: $M_{\text{Targets}} = 648$, $SE_{\text{Targets}} = 23$, $M_{\text{Nontargets}} = 578$, $SE_{\text{Nontargets}} = 19$; Valid-cue condition: $M_{\text{Targets}} = 518$, $SE_{\text{Targets}} = 16$, $M_{\text{Nontargets}} = 504$, $SE_{\text{Nontargets}} = 16$, $t(20) = 8.53$, $p < .001$).

In addition, the fact that rather high accuracy rates were observed within the two *different* conditions of the target trials (averaged 0.86 across experiments; see Table 2 below) further indicated that participants obeyed task instructions and responded to the cued object only, as distractor stimuli in these conditions were associated with an opposing response. Table 2 presents the results obtained for the different object-pair conditions in Experiments 2A and 2B.

An inspection of the categorical relation effect among the nontarget trials revealed no effect of categorical similarity with the RT measure [Exp. 2A: $t(20) = 0.01$, $p > .1$, $d = 0.01$, BF = 0.23; Exp. 2B: $t(20) = 0.31$, $p > .1$, $d = 0.07$, BF = 0.24] or the accuracy rate measure [Exp. 2A: $t(20) = 0.35$, $p > .1$, $d = 0.08$, BF = 0.24; Exp. 2B: $t(20) = 1.51$, $p > .1$, $d = 0.33$, BF = 0.61, which is inconclusive according to Bayesian

Table 2 Mean reaction times (in milliseconds) and accuracy rates for trials containing pair objects within the categorical classification task of Experiments 2A and 2B (standard errors are in parentheses)

	Nontarget Trials		Target Trials		
	Same	Different	Same	Different	
				Target Cued	Target Uncued
Experiment 2A (Target: meaningless shape)					
Reaction time	501 (12)	500 (13)	532 (14)	535 (14)	535 (16)
Accuracy rates	.9 (.01)	.9 (.01)	.83 (.02)	.84 (.02)	.88 (.02)
Experiment 2B (Target: real-world category)					
Reaction time	503 (15)	504 (15)	521 (14)	524 (15)	553 (16)
Accuracy rates	.91 (.01)	.91 (.02)	.87 (.01)	.85 (.02)	.86 (.02)

norms], suggesting that object-to-object associative and visual processing was absent when one of the two stimuli was

positioned at an unattended location (see Fig. 5). To further explore the differences between Experiment 1 (both image

Fig. 5 The *categorical relation effect* and the *target distraction effect* in Experiments 2A and 2B. Bars represent the group average difference within each contrast, and circles represent the individual participants' difference scores. Black horizontal lines denote one standard error above and below the group average, computed for the difference response time (RT) scores (e.g., Loftus & Masson, 1994). The image examples underneath the bars represent the specific contrast computed, with black frames denoting the cue appearing prior to stimulus presentation. The cue appeared with equal probability in the upper and lower locations, and it disappeared once the images appeared (it is presented here for illustrative purposes). The categorical relation effect was computed as the difference between responses to different-category displays (e.g., a cued vehicle and an uncued animal) and

same-category displays (e.g., two types of vehicles, cued and uncued), within the nontarget trials. The target distraction effect was computed as the RT difference between the different-category condition within the target trials (*target-uncued*, left) and the different-category condition within the nontarget trials (right). Note that in both cases, the categories in the cued and uncued locations differed from each other, yet only in the former condition did a to-be-detected stimulus appear outside focal attention. This distraction caused by a prioritized target positioned in an unattended location was particularly important in Experiment 2B, in which real-world objects served as the targets, and thus any obvious perceptual differences between the two different-category conditions were greatly reduced. See the main text for further explanation and discussion

locations were relevant to the task requirements) and Experiment 2 (one of two image locations was task-irrelevant), a two-way mixed analysis of variance (ANOVA) was conducted on the RTs of the nontarget trials, including categorical relation (same, different) as a within-subjects factor, and experiment (1, 2) as a between-subjects factor. When using meaningless shapes as the target items (Exp. 1A, 2A), this analysis with RTs revealed a significant experimental main effect [$F(1, 40) = 5.04, p = .03, \eta_p^2 = .11, BF = 1.26$], as well as a categorical relation effect [$F(1, 40) = 6.25, p = .017, \eta_p^2 = .14, BF = 1.4$] that was qualified by a significant interaction effect [$F(1, 40) = 6.64, p = .014, \eta_p^2 = .14, BF = 3.83$]. When using real-world objects as the target category (Exp. 1B, 2B), the corresponding analysis yielded a categorical main effect [$F(1, 40) = 5.95, p < .02, \eta_p^2 = .13, BF = 2.9$], qualified by a (marginally) significant interaction effect [$F(1, 40) = 3.83, p = .057, \eta_p^2 = .09, BF = 0.73$]. The experimental factor was nonsignificant ($F < 1, p > .1, \eta_p^2 = .02, BF = 1.01$). Although the interaction effect was mild (and inconclusive, according to Bayesian norms), these results imply that the insertion of the spatial-cuing manipulation in Experiment 2B attenuated object-to-object categorical processing (effectively nullifying them). When I conducted a two-way mixed ANOVA with the accuracy rate measure, no statistically significant effects were observed across both experiments (all F s < 2 , all p s $> .1$).

Taken together, the RT findings of both sets of studies portray a rather clear picture, according to which linking two objects on the basis of their visual and/or categorical identity involves spatial attentional allocation. Participants exhibited a sensitivity to object visual/categorical relations if both items appeared in an attended location (Exp. 1). These categorical effects, however, largely disappeared when one of two items was presented in an unattended, or, a minimally attended location (Exp. 2).

Interestingly, an uncued stimulus *did* affect task performance when it contained a to-be-detected target item that differed from the cued (nontarget) item. That is, when contrasting the target-uncued (different-category) condition with the nontarget, different-category condition, a robust difference in RT performance was observed [Exp. 2A: $t(20) = 5.64, p < .001, d = 1.23, BF > 100$; Exp. 2B: $t(20) = 10.78, p < .001, d = 2.4, BF > 100$]. Note that both conditions contained a cued nontarget stimulus, paired with an uncued image from a different category. Yet only the target-uncued condition contained a stimulus from a prioritized category in the distractor's location, causing a robust distraction effect (see the "Target distraction" bars in Fig. 5). Although this distraction could be potentially accounted for by the visual salience of the meaningless shapes used in Experiment 2A, perceptual salience cannot explain the results of Experiment 2B, in which real-world objects were used as the target categories (counterbalanced across category type). The present findings suggest that processing the

relations of items taken from a prioritized category with items from other, nonprioritized categories persists even when the former are positioned at the outskirts of focal attention (see the further elaboration on this point in Exp. 4).

And yet, recall that the focus of the study was on everyday situations, in which an attended object is surrounded by items that belong to the same superordinate category, but are not necessarily exemplars of the same basic-level category (e.g., a sofa lying near a couch or a table). As was mentioned in the introduction, one of the factors that may potentially support unattended categorical processing is the set of visual features shared among same-category objects. It was reasoned that a potential visual overlap among items may reduce perceptual ambiguity among stimuli appearing outside the focus of attention, possibly allowing their partial recognition despite the underprivileged visual conditions. In Experiments 1 and 2, obvious repetitiveness of individual exemplar images within each category was avoided, in order to prevent participants from actively tuning themselves to highly specific visual features or object configurations. Yet, perhaps the use of wide superordinate categories induced large visual variability that did not allow sufficient predictability outside focal attention. The use of narrower, more homogeneous categories, would potentially enable activation of visually similar items on the one hand, and allow a better distinction between visually dissimilar categories, when appearing in unattended locations. That is, although no categorical effects were observed among the nontarget distractor stimuli when using superordinate categories, perhaps the use of basic-level ones would allow higher visual predictability and a reliable categorical effect with such stimuli.

To test whether categorical heterogeneity affects object-object relational processing, two additional studies were conducted. In the first, three highly distinct real-world, basic-level categories were used; in the second, three basic-level categories were presented, however, they now overlapped in their semantic and visual properties. That is, the three basic-level categories were taken from the same superordinate category, thus they shared many semantic as well as visual characteristics. It was hypothesized that to the extent that reducing the within-category heterogeneity increases perceptual predictability outside focal attention, a categorical relation effect may be observed even when one of two objects appears in an unattended location. Importantly, the more visually distinctive the categories, the finer the representation of each of the categories as well as the representation of the relations among them (i.e., whether two objects belong to the same or to different categories). When using the spatial-cuing manipulation, then, a categorical relation effect may be observed among the nontarget stimuli in the first experiment (highly distinct categories), but not in the second one (highly overlapping categories), due to a poor representation of the relations among the categories in the latter.

Experiments 3A and 3B: Visual and categorical relations among basic-level categories (one of two objects is an irrelevant distractor)

Method

Participants Twenty-one undergraduate students (16 females, five males; ages 21–36 years) participated in Experiment 3A, and an additional independent group of 21 students (15 females, six males; ages 19–34 years) participated in Experiment 3B. All participants had normal or corrected-to-normal sight and received course credit for participation in the experiments.

Stimuli, apparatus, and experimental design Both experiments were carried out using the spatial-cuing paradigm, in which one of two objects was cued (attended) while its counterpart object appeared in an uncued location and was irrelevant to task requirements (hence, “unattended,” or minimally attended). The three highly distinct categories used in Experiment 3A were chairs, cars and houses, whereas the three highly overlapping categories used in Experiment 3B were chairs, chests and tables (all are types of furniture). All other aspects of the experiments were identical to those in Experiment 2.

Results and discussion

Mean RTs and accuracy rates were computed for each condition, within each experiment, see Table 3. Trials on which participants made errors (Exp. 3A, 13%; Exp. 3B, 15%) or trials on which there were extreme outlier responses (approximately 1.5%, 1%, respectively) were excluded from the RT analysis. As in the previous experiments, the cuing manipulation was effective in both experiments, as evident from the cue-validity latency effect [Exp. 3A: Invalid-cue condition:

$M_{\text{Targets}} = 641$, $SE_{\text{Targets}} = 28$, $M_{\text{Nontargets}} = 540$, $SE_{\text{Nontargets}} = 23$; Valid-cue condition: $M_{\text{Targets}} = 514$, $SE_{\text{Targets}} = 15$, $M_{\text{Nontargets}} = 488$, $SE_{\text{Nontargets}} = 16$, $t(20) = 5.21$, $p < .001$, averaged across target and nontarget items; Exp. 3B: Invalid: $M_{\text{Targets}} = 667$, $SE_{\text{Targets}} = 28$, $M_{\text{Nontargets}} = 600$, $SE_{\text{Nontargets}} = 22$; Valid: $M_{\text{Targets}} = 551$, $SE_{\text{Targets}} = 18$, $M_{\text{Nontargets}} = 537$, $SE_{\text{Nontargets}} = 15$, $t(20) = 5.52$, $p < .001$].

The relatively high accuracy rates obtained in the two *different* conditions of the target trials within these experiments (averaged .83; see Table 3) further suggest that participants obeyed the task instructions and responded to the cued location.

A small-to-medium, yet statistically significant, categorical relation effect was obtained with the RT measure when using three highly distinct basic-level categories [Exp. 3A: $t(20) = 2.18$, $p = .042$, $d = 0.48$, $BF = 1.58$]. This effect was nonsignificant when the three categories were taken from the same superordinate category and contained overlapping features [Exp. 3B: $t(20) = 1.2$, $p > .1$, $d = 0.26$, $BF = 0.43$]. A direct contrast between the results of the two experiments revealed no significant difference [$F(1, 40) < 1$, $p > .1$, $\eta_p^2 = .01$, $BF = 0.36$; an inconclusive finding according to Bayesian norms, as were those in the two previous analyses]. None of the contrasts with the arcsine-transformed mean accuracy rates were statistically significant [Exp. 3A: $t(20) = 1.81$, $p > .05$, $d = 0.4$, $BF = 0.9$; Exp. 3B: $t(20) = 0.01$, $p > .1$, $d = 0.02$, $BF = 0.23$].

In contrast, a clear target distraction effect (computed as the target-uncued vs. the different-category nontarget condition) was observed in both experiments (see Fig. 6). Once again, a strong interference was caused by a to-be-detected target category positioned in a distractor’s location [Exp. 3A: $t(20) = 6.75$, $p < .001$, $d = 1.47$, $BF > 100$; Exp. 3B: $t(20) = 5.15$, $p < .001$, $d = 1.22$, $BF > 100$]. There was no significant difference in magnitude between the distraction effects of the two experiments [$F(1, 40) = 1.4$, $p > .1$, $\eta_p^2 = .02$, $BF = 0.54$]. A robust target distraction effect was obtained also with the arcsine-transformed accuracy rates in Experiment 3A [$t(20) = 3.77$,

Table 3 Mean reaction times (in milliseconds) and accuracy rates in the categorical classification task, among the pair object trials of Experiments 3A and 3B (standard errors are in parentheses)

	Nontarget Trials		Target Trials		
	Same	Different	Same	Different	
				Target Cued	Target Uncued
Experiment 3A (chairs, cars, houses)					
Reaction time	483 (17)	490 (17)	519 (16)	525 (15)	545 (20)
Accuracy rates	.93 (.01)	.91 (.01)	.84 (.02)	.80 (.03)	.85 (.02)
Experiment 3B (chairs, chests, tables)					
Reaction time	536 (14)	540 (13)	556 (15)	554 (15)	580 (16)
Accuracy rates	.88 (.02)	.88 (.02)	.82 (.02)	.83 (.02)	.83 (.04)

Fig. 6 The *categorical relation* effect and the *target distraction* effect in Experiments 3A and 3B. The categorical relation effect was computed as the difference between different-category displays (e.g., a cued car and an uncued house) and same-category displays (e.g., two cars, cued and uncued), within the nontarget trials. The target distraction effect was

computed as the response time (RT) difference between the different-category condition within the target trials (*target-uncued*, left) and the different-category condition within the nontarget trials (right). The target category in the present example is *Chairs*. For further details, see Fig. 5

$p = .001$, $d = 0.82$, $BF = 30.7$], but not in Experiment 3B [$t(20) = 1.19$, $p > .1$, $d = 0.26$, $BF = 0.42$]. There was no significant difference among the results of the two experiments [$F(1, 40) = 1.35$, $p > .1$, $\eta_p^2 = .01$, $BF = 0.49$].

Taken together, the manipulation of within-category homogeneity along with the between-category distinctiveness affected the processing of object-to-object visual and categorical relations, yet only to a mild degree. That is, using narrowed, basic-level categories that were highly distinct from each other allowed some associative processing when one of two items was presented outside the main focus of attention. As was mentioned above, however, these effects were moderate in size, and they were completely eliminated when objects were characterized by a relatively large between-category visual overlap. This latter result was expected, since any benefit of a within-category feature repetitiveness (potentially reducing perceptual uncertainty in unattended locations) was cancelled out by the difficulty to dissociate the three different categories. As in Experiment 2, however, a consistent RT distraction effect was observed, elicited by a prioritized object category appearing at an irrelevant distractor's location. That is, a target object appearing outside focal attention and differing from the category of an attended object affected performance dramatically regardless of category nature (i.e., the degree of its

homogeneity or its overall distinctiveness). This finding can be attributed to the powerful impact that unattended stimuli have on behavior when complying with the current task settings (e.g., when they are actively searched for; see, e.g., Bacon & Egeth, 1994; Folk et al., 1992).

Note, however, that in all experiments the distraction effect included a strong response conflict component, similar to the one characterizing typical flanker tasks. That is, in contrast to the nontarget *different* condition, the to-be-detected target category appearing outside focal attention (i.e., in the target-uncued condition) not only differed from the cued object in its category, but also in its associated response (“target,” in contrast to the cued “nontarget” response). This conflict, and the potential requirement to inhibit the activated (but unwanted) response caused by the distractor may have slowed down responses in the target-uncued condition. It is possible, then, that the distraction effect was accounted for by response conflict factors, rather than by pure visual and categorical similarity factors.

To control for response conflict factors among the target category stimuli, a fourth experiment was conducted, in which object-to-object relations were now assessed while holding constant the response activation factor across cued and uncued locations. To this end, *two* target categories were chosen in

addition to the two nontarget categories. The use of two target categories allowed us to examine the “pure” categorical relation effect within the target trials, and to compare it to the respective effect within the nontarget trials (see below).

Experiment 4: Further investigating the categorical relation effect among prioritized (target) categories

Method

Participants Twenty-four³ undergraduate students (13 females, 11 males; ages 19–40 years) participated in Experiment 4. All participants had normal or corrected-to-normal sight and received course credit for participation in the experiments.

Stimuli, apparatus, and experimental design The experiment was carried out using the spatial cuing paradigm, similar to the one used in previous experiments. To the three superordinate categories used in Experiments 1 and 2 (i.e., animals, vehicles, and clothes), a fourth category of furniture was added. As was mentioned above, two of the categories were defined as target categories (e.g., animals, vehicles), whereas the other two were nontarget categories (e.g., clothes, furniture; targets and nontargets counterbalanced across participants). As in the previous experiments, among trials containing nontarget pair stimuli (240 trials), half of the pairs contained same-category objects, whereas the other half contained different-category objects. In the present experiment, however, a similar division existed among the trials containing *two* target stimuli (240 trials): Half were taken from the same category (e.g., both clothing articles), whereas the other half were taken from different target categories (e.g., an article of clothing and a piece of furniture). Note that the latter, *different-category* condition allowed for assessing categorical relations within the target-present trials, while holding constant the response factor within the cued and the uncued locations, as the distractor stimuli in the same- and different-category target conditions were both associated with a “target” response.

To further assess the target distraction effect, in which a to-be-detected target positioned in a distractor location interfered (and conflicted) with the response to a cued, *nontarget* stimulus, the “target-uncued” condition used in previous experiments, was maintained (120 trials; e.g., an uncued piece of furniture paired with a cued vehicle). The “target-cued” condition, as well as all single-item conditions (previously used for the assessment of cue-validity effects) were omitted from

the experiment. All other aspects of the experiment were identical to those in Experiment 3.

Results and discussion

Mean RTs and accuracy rates were computed for each condition. Trials on which participants made errors (13.3%) or trials on which there were extreme outlier responses (approximately 1.5%) were excluded from the RT analysis. Table 4 presents the results obtained for the different object-pair conditions in Experiment 4. A two-way repeated measures ANOVA was conducted on the response latency data, including categorical relation (same, different) and target presence (nontarget, target) as factors. The ANOVA revealed no target presence effect [$F(1, 23) = 0.09, p > .1, \eta_p^2 = .004, BF = 0.2$], yet a significant categorical relation effect [$F(1, 23) = 8.88, p < .007, \eta_p^2 = .28, BF = 0.66$] that was qualified by an interaction [$F(1, 23) = 5.31, p < .031, \eta_p^2 = .19, BF = 0.69$]. Although the interaction was inconclusive according to Bayesian norms, follow-up contrasts clearly revealed its origin: a strong categorical effect was observed among the target stimuli [$t(23) = 4.14, p < .001, d = 0.84, BF = 78$], whereas there was no hint for such an effect among the nontarget trials [$t(23) = -0.01, p > .1, d = 0.002, BF = 0.22$]. In addition, and in accord with previous findings, a robust target distraction effect was observed [$t(23) = 4.74, p < .001, d = 0.97, BF > 100$] (see Fig. 7).

A similar two-way ANOVA conducted on the arcsine-transformed mean accuracy rates of Experiment 4 revealed a significant target presence effect [$F(1, 23) = 13.43, p < .001, \eta_p^2 = .37; BF > 100$], yet no categorical relation effect [$F(1, 23) = 2.02, p > .1, \eta_p^2 = .08; BF = 0.30$] or interaction ($F < 1, BF = 0.29$). None of the follow-up contrasts were significant ($ts \leq 1, 0.32 < BFs < 0.35$). The target distraction effect was again significant [$t(23) = 3.79, p < .001, d = 0.77, BF = 36$].

The results of Experiment 4 converged on the previous results (when using superordinate categories) in showing no categorical relation effect among the nontarget trials, when one of two stimuli appeared outside the main focus of spatial attention. Interestingly, however, a significant categorical effect was observed among the *target* trials, indicating that response conflict and/or response inhibition are not necessary factors for target distraction effects on behavior. That is, in contrast to nonprioritized (i.e., nontarget) stimuli, the visual and categorical relations among target stimuli are clearly processed and may influence behavior even when one of two stimuli is presented in an unattended location.

General discussion

The present study investigated the involvement of visual attention in object-to-object categorical processing during a brief visual glimpse. This issue has attracted growing attention

³ To allow for full counterbalancing of all possible combinations of pair targets, a sample of 24 participants was chosen.

Table 4 Mean reaction times (in milliseconds) and accuracy rates in the different conditions of Experiment 4 (standard errors are in parentheses)

	Nontarget Trials		Target Trials		
	<i>Same</i>	<i>Different</i>	<i>Same</i>	<i>Different</i>	<i>Target Uncued</i>
Reaction time	573 (14)	573 (13)	569 (14)	580 (14)	597 (14)
Accuracy rates	.90 (.08)	.89 (.08)	.84 (.09)	.84 (.08)	.86 (.09)

ever since pioneering studies investigating the nature of attentional mechanisms and their effects on object recognition. Previous attentional studies have examined attentional necessity to individual stimulus processing, as well as to more global object-to-object relational processing. The present study used a focused attention paradigm, in which responses to stimuli were assessed when visual attention spanned both objects versus when attention was focused on only one object while its counterpart pair object served as a task-irrelevant distractor. Objects in each pair were either taken from the same or from different superordinate categories. The results showed that when participants attended to both objects, a categorical relation effect was obtained (Exp. 1). Namely, participants were sensitive to whether nonprioritized (i.e., nonsearched) object pairs belonged to same or to different superordinate categories. Focusing attention on only one of two stimuli, however,

largely eliminated this effect (Exp. 2), indicating that visual attention has an important role in object-to-object visual and categorical processing. Nevertheless, a moderate categorical relation effect was evident under the latter conditions when categories were highly distinct and were narrowed to basic level. This finding implies that visual and categorical relations among objects may be registered if stimuli share (or differ by) many visual features, hence, the unattended input is highly predictable (Exp. 3). Critically, when a prioritized (i.e., to-be-detected) object category, positioned outside the main focus of attention, differed from an attended object, a robust distraction was observed even when using superordinate categories. This effect did not rely on response conflict or response inhibition factors (Exp. 4). These latter findings suggest that object relations that involve stimuli that are relevant to goal-directed settings may survive attentional deprivation in

Fig. 7 The *categorical relation* effect within each of the target presence conditions, as well as the *target distraction* effect in Experiment 4. The two target categories in the present example are *Clothes* and *Furniture*.

Note that the categorical relation effect was now assessed under both nontarget and target conditions. Response activation was held constant among the target trials, similarly to the nontarget trials

the unattended location, in spite of a large inter-categorical heterogeneity. The involvement of spatial attention in superordinate categorical (and visual) relational processing is most critical in situations that include stimuli that are *not* prioritized by current goals in any way.

Relation to previous findings

The question of the involvement of attention in object-to-object relational processing is fundamental to understanding the breadth, and the constraints, of scene perception “within a mere glimpse”. The focus of the present study is on everyday situations, in which an attended object is surrounded by items that share an overall context (i.e., superordinate category), but do not necessarily share the same basic-level category. A couch in a living room may lay near a sofa or a table, and a truck on the road may be seen near a car or a motorcycle. As was mentioned in the introduction, previous research has shown that under unattended, or minimally attended conditions, the processing of the semantic relations between individual objects is greatly reduced (Gronau & Shahar, 2014; Malcolm et al., 2016). Objects belonging to the same superordinate category, however, do not only belong to the same overall semantic world but often share some visual features that may assist in disambiguating object identity in the unattended location. Indeed, ample research has emphasized the importance of visual similarity factors in the processing of stimulus relations and their effects on item detection and/or identification. For example, paradigms studying the redundancy gain phenomenon or studies using visual search tasks have examined the effects of visual similarity (among targets, or targets and distractors, respectively) on target detection. In a similar vein, studies using the flanker task have manipulated target–distractor similarity and its influence on the flanker distraction effect. Note that most previous studies have used rather simple, arbitrary/symbolic stimuli that do not represent the rich variety of real-world objects belonging to the same superordinate category. More importantly, the stimuli in such studies were typically attended (or, partially attended), and/or were highly relevant to one’s task settings. Thus, for instance, in the visual search paradigm, distractor stimuli are typically brought into the main focus of spatial attention via active search. In addition, guidance of attention is mediated by featural and perhaps conceptual properties of the searched target (Seidl-Rathkopf et al., 2015; Wolfe et al., 1989; see the review in Wolfe & Horowitz, 2017). Target–distractor similarity effects, thus, may be influenced to a large extent by goal-driven attentional settings. Studies using the flanker task, in contrast, adopt a focused-attention approach (in which flankers are positioned at the outskirts of focal attention), yet target and distractors are normally taken from a limited, predefined set of to-be-detected stimuli. As was noted earlier, the distinction between prioritized (to-be-detected, or

searched) and non-prioritized (nonsearched) stimuli is crucial when examining the effects of irrelevant distractors presented at unattended locations, since the former type of items are known to exert a strong influence on performance via feature-based, conceptual-based, or other goal-directed attentional mechanisms (Bacon & Egeth, 1994; Folk et al., 1992; Hillyard & Münte, 1984; Müller et al., 1995; Wyble et al., 2013). Indeed, the present study corroborates these findings, in demonstrating a robust target distraction effect (Exps. 2–4) and a categorical relation effect among items belonging to the target category (Exp. 4). These effects emphasize the strong influence of unattended distractors, prioritized by online task settings, on attended objects. In contrast, the present findings reveal that when an object does *not* belong to a prioritized superordinate category, the processing of its visual and/or semantic relations with an attended object is largely eliminated.

Importantly, the question of the involvement of visual attention in processing objects’ interrelations taps into more basic questions concerning the necessity of spatial attention to processing the building blocks of these relations—that is, the individual objects. In fact, such questions are at the heart of the historical “early-versus-late” selection debate (e.g., Broadbent, 1958; Deutsch & Deutsch, 1963; Duncan, 1980; Treisman, 1960). Although this debate subsided during the 90s of the last century (partially due to the perceptual load theory proposed by Lavie, 1995), a new wave of studies in recent years has brought the dispute back to center stage. In contrast to most early research, these studies have used real-world objects (rather than arbitrary or symbolic stimuli), arguing that such objects benefit from extensive training and high familiarity levels, and thus they form a unique class of stimuli that are detected and recognized even outside the main focus of attention (e.g., Li, VanRullen, Koch, & Perona, 2002; Otsuka & Kawaguchi, 2007; Poncet, Reddy, & Fabre-Thorpe, 2012; Reddy, Reddy, & Koch, 2006). Importantly, the bulk of the studies arguing for object categorization and identification “in the near absence of attention” have used dual tasks in which objects’ identities were, in effect, highly prioritized by the task settings (e.g., Li et al., 2002; Poncet et al., 2012; VanRullen, Reddy, & Fei-Fei, 2005). Namely, participants were actively instructed to detect an object category (e.g., an animal appearing in the visual periphery), while performing a simultaneous secondary task (search at fixation). There are good reasons to believe that while performing the dual task, participants spread their attention to areas in which target objects could potentially appear. Furthermore, the prioritization of the target objects by task demands most likely activated feature-based and/or other goal-directed attentional mechanisms that assisted in target detection and identification, despite the appearance of these objects in an underprivileged location. As is evident from the present study, dissociating “prioritized” and “nonprioritized” stimuli can be crucial when discussing object processing under unattended conditions.

Indeed, several follow-up studies failed to replicate the findings of attention-free image categorization and recognition, casting doubt on the notion of unlimited capacity for object and scene recognition (e.g., M. A. Cohen, Alvarez, & Nakayama, 2011; Evans & Treisman, 2005; Gronau & Izoutcheev, 2017; Mack & Clarke, 2012; Potter & Fox, 2009; Scharff, Palmer, & Moore, 2011; Walker, Stafford, & Davis, 2008).

In conclusion, the present study focused on the involvement of spatial attention in object-to-object visual and categorical relations. The findings of the study suggest that processing the relations between attended and unattended objects relies on the relevance of the latter to the task requirements. Objects belonging to a prioritized superordinate category exert strong influence on other, nonprioritized stimuli, despite their appearance outside the main focus of visual attention. Unattended items that are not prioritized by the task demands, however, do not exert influence on attended items, unless the objects' categories are narrowed to the basic level and are highly distinct from each other. Although the unique status of stimuli possessing task-relevant information has long been documented, it is yet to be determined whether the categorization of such stimuli (positioned in unattended locations) is based on coarse detection of visual features/object parts (e.g., Evans & Treisman, 2005; Levin, Takarae, Miner, & Keil, 2001; Walther & Shen, 2014) or on a more abstract, conceptual representation of the target category (e.g., Peelen et al., 2009; Seidl-Rathkopf et al., 2015; Wyble et al., 2013; see also VanRullen, 2009). Future research will directly examine these questions.

Author Note I thank S. Yakim, A. Izoutcheev, M. Shachar, and T. Nave for their assistance in experiment programming and data collection. This research was supported by the Israel Science Foundation (grant no. 1622/15) to N.G.

References

- Antes, J. R., Penland, J. G., & Metzger, R. L. (1981). Processing global information in briefly presented pictures. *Psychological Research*, 43, 277–292.
- Auckland, M. E., Cave, K. R., & Donnelly, N. (2007). Nontarget object can influence perceptual processes during object recognition. *Psychonomic Bulletin & Review*, 14, 332–337. <https://doi.org/10.3758/BF03194073>
- Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, 55, 485–496. <https://doi.org/10.3758/BF03205306>
- Bajo, M.-T. (1988). Semantic facilitation with pictures and words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 579–589. <https://doi.org/10.1037/0278-7393.14.4.579>
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, 5, 617–629. <https://doi.org/10.1038/nrn1476>
- Biederman, I. (1972). Perceiving real-world scenes. *Science*, 177, 77–80. <https://doi.org/10.1126/science.177.4043.77>
- Biederman, I. (1981). Do background depth gradients facilitate object identification? *Perception*, 10, 573–578. <https://doi.org/10.1068/p100573>
- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, 14, 143–177. [https://doi.org/10.1016/0010-0285\(82\)90007-X](https://doi.org/10.1016/0010-0285(82)90007-X)
- Boyce, S. J., Pollatsek, A., & Rayner, K. (1989). Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 556–566.
- Broadbent, D. E. (1958). *Perception and communication*. Elmsford, NY, US: Pergamon Press. <https://doi.org/10.1037/10037-000>
- Carr, T. H., McCauley, C., Sperber, R. D., & Parnellee, C. M. (1982). Words, pictures, and priming: On semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 757–777. <https://doi.org/10.1037/0096-1523.8.6.757>
- Clarke, A., & Tyler, L. K. (2015). Understanding what we see: How we derive meaning from vision. *Trends in Cognitive Sciences*, 19, 677–687.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological Science*, 22, 1165–1172. <https://doi.org/10.1177/0956797611419168>
- Davenport, J. L. (2007). Consistency effects between objects in scenes. *Memory & Cognition*, 35, 393–401. <https://doi.org/10.3758/BF03193280>
- Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, 15, 559–564. <https://doi.org/10.1111/j.0956-7976.2004.00719.x>
- Demiral, Ş. B., Malcolm, G. L., & Henderson, J. M. (2012). ERP correlates of spatially incongruent object identification during scene viewing: Contextual expectancy versus simultaneous processing. *Neuropsychologia*, 50, 1271–1285.
- Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, 70, 80–90. <https://doi.org/10.1037/h0039515>
- Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, 6, 274–290. <https://doi.org/10.1177/1745691611406920>
- Dobel, C., Gummior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, 125, 129–143. <https://doi.org/10.1016/j.actpsy.2006.07.004>
- Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, 87, 272–300. <https://doi.org/10.1037/0033-295X.87.3.272>
- Duncan, J., & Humphreys, G. W. (1989). Visual-search and stimulus similarity. *Psychological Review*, 96, 433–458. <https://doi.org/10.1037/0033-295X.96.3.433>
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 143–149. <https://doi.org/10.3758/BF03203267>
- Eriksen, C. W., & Eriksen, B. A. (1979). Target redundancy in visual search: Do repetitions of the target within the display impair processing? *Perception & Psychophysics*, 26, 195–205.
- Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance*, 31, 1476–1492. <https://doi.org/10.1037/0096-1523.31.6.1476>
- Fabre-Thorpe, M. (2011). The characteristics and limits of rapid visual categorization. *Frontiers in Psychology*, 2, 243. <https://doi.org/10.3389/fpsyg.2011.00243>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral,

- and biomedical sciences. *Behavior Research Methods*, 39, 175–191. <https://doi.org/10.3758/BF03193146>
- Folk, C. L., & Remington, R. (1998). Selectivity in attentional capture by featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 847–858. <https://doi.org/10.1037/0096-1523.24.3.847>
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 1030–1044. <https://doi.org/10.1037/0096-1523.18.4.1030>
- Fournier, L., Scheffers, M. K., Coles, M. G., Adamson, A., & Abad, E. V. (1997). The dimensionality of the flanker compatibility effect: A psychophysiological analysis. *Psychological Research*, 60, 144–155.
- Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, 108, 316–355.
- Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research*, 16, 123–144.
- Green, C., & Hummel, J. E. (2006). Familiar interacting object pairs are perceptually grouped. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1107–1119. <https://doi.org/10.1037/0096-1523.32.5.1107>
- Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science*, 16, 152–160. <https://doi.org/10.1111/j.0956-7976.2005.00796.x>
- Gronau, N., Cohen, A., & Ben-Shachar, G. (2003). Dissociations of personally significant and task-relevant distractors inside and outside the focus of attention: A combined behavioral and psychophysiological study. *Journal of Experimental Psychology: General*, 132, 512–529. <https://doi.org/10.1037/0096-3445.132.4.512>
- Gronau, N., Cohen, A., & Ben-Shachar, G. (2009). Distractor interference in focused attention tasks is not mediated by attention capture. *Quarterly Journal of Experimental Psychology*, 62, 1685–1695. <https://doi.org/10.1080/17470210902811223>
- Gronau, N., & Izoutcheev, A. (2017). The necessity of visual attention to scene categorization: Dissociating “task-relevant” and “task-irrelevant” scene distractors. *Journal of Experimental Psychology: Human Perception and Performance*, 43, 954–970.
- Gronau, N., & Shachar, M. (2014). Contextual integration of visual objects necessitates attention. *Attention, Perception, & Psychophysics*, 76, 695–714. <https://doi.org/10.3758/s13414-013-0617-8>
- Henderson, J. M., Pollatsek, A., & Rayner, K. (1987). Effects of foveal priming and extrafoveal preview on object identification. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 449–463. <https://doi.org/10.1037/0096-1523.13.3.449>
- Hillyard, S. A., & Münte, T. F. (1984). Selective attention to color and location: An analysis with event-related brain potentials. *Perception & Psychophysics*, 36, 185–198. <https://doi.org/10.3758/BF03202679>
- Hollingworth, A. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General*, 127, 398–415. <https://doi.org/10.1037/0096-3445.127.4.398>
- Holmgren, J. E., Juola, J. F., & Atkinson, R. C. (1974). Response latency in visual search with redundancy in the visual display. *Perception & Psychophysics*, 16, 123–128.
- Hogg, R. V., & Craig, A. T. (1995). Introduction to mathematical statistics (5th ed.). Englewood Hills, NJ: Wiley.
- Jacobs, A. M. (1986). Eye-movement control in visual search: How direct is visual span control? *Perception & Psychophysics*, 39, 47–58.
- Jeffreys, H. (1961). Theory of probability. Oxford, UK: Oxford University Press.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175–219. [https://doi.org/10.1016/0010-0285\(92\)90007-O](https://doi.org/10.1016/0010-0285(92)90007-O)
- Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning and Verbal Behavior*, 23, 39–66.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 451–468. <https://doi.org/10.1037/0096-1523.21.3.451>
- Levin, D. T., Takarae, Y., Miner, A. G., & Keil, F. (2001). Efficient visual search by category: Specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception & Psychophysics*, 63, 676–697. <https://doi.org/10.3758/BF03194429>
- Li, F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, 99, 9596–9601.
- Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1, 476–490. <https://doi.org/10.3758/BF03210951>
- Mack, A., & Clarke, J. (2012). Gist perception requires attention. *Visual Cognition*, 20, 300–327. <https://doi.org/10.1080/13506285.2012.666578>
- Mack, A., & Rock, I. (1998). Inattention blindness. Cambridge, MA, US: MIT Press.
- Malcolm, G. L., Rattinger, M., & Shomstein, S. (2016). Intrusive effects of semantic information on visual selective attention. *Attention, Perception, & Psychophysics*, 78, 2066–2078.
- Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, 29, 317–322. <https://doi.org/10.1016/j.tins.2006.04.001>
- Memelink, J., & Hommel, B. (2013). Intentional weighting: A basic principle in cognitive control. *Psychological Research*, 77, 249–259.
- Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 520–538. <https://doi.org/10.1037/0096-1523.17.2.520>
- Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics*, 57, 1–17. <https://doi.org/10.3758/BF03211845>
- Müller, H. J., Reimann, B., & Krummenacher, J. (2003). Visual search for singleton feature targets across dimensions: Stimulus- and expectancy-driven effects in dimensional weighing. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 1021–1035. <https://doi.org/10.1037/0096-1523.29.5.1021>
- Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, 11, 520–527. <https://doi.org/10.1016/j.tics.2007.09.009>
- Oppermann, F., Hassler, U., Jescheniak, J., & Gruber, T. (2012). The rapid extraction of gist—Early neural correlates of high-level visual processing. *Journal of Cognitive Neuroscience*, 24, 521–529. https://doi.org/10.1162/jocn_a_00100
- Otsuka, S., & Kawaguchi, J. (2007). Natural scene categorization with minimal attention: Evidence from negative priming. *Perception & Psychophysics*, 69, 1126–1139.
- Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, 3, 519–526. <https://doi.org/10.3758/BF03197524>
- Peelen, M. V., Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature*, 460, 94–97. <https://doi.org/10.1038/nature08103>
- Poncet, M., Reddy, L., & Fabre-Thorpe, M. (2012). A need for more information uptake but not focused attention to access basic-level

- representations. *Journal of Vision*, 12(1), 15. <https://doi.org/10.1167/12.1.15>
- Potter, M. C., & Fox, L. F. (2009). Detecting and remembering simultaneous pictures in a rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 28–38. <https://doi.org/10.1037/a0013624>
- Reddy, L., Reddy, L., & Koch, C. (2006). Face identification in the near-absence of focal attention. *Vision Research*, 46, 2336–2343. <https://doi.org/10.1016/j.visres.2006.01.020>
- Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8, 368–373. <https://doi.org/10.1111/j.1467-9280.1997.tb00427.x>
- Roberts, K. L., & Humphreys, G. W. (2011). Action relations facilitate the identification of briefly-presented objects. *Attention, Perception, & Psychophysics*, 73, 597–612. <https://doi.org/10.3758/s13414-010-0043-0>
- Roper, Z. J. J., Cosman, J. D., & Vecera, S. P. (2013). Perceptual load corresponds with factors known to influence visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 39, 1340–1351. <https://doi.org/10.1037/a0031616>
- Scharff, A., Palmer, J., & Moore, C. M. (2011). Evidence of fixed capacity in visual object categorization. *Psychonomic Bulletin & Review*, 18, 713–721. <https://doi.org/10.3758/s13423-011-0101-1>
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime reference guide. Pittsburgh, PA: Psychology Software Tools, Inc.
- Seidl-Rathkopf, K. N., Turk-Browne, N. B., & Kastner, S. (2015). Automatic guidance of attention during real-world visual search. *Attention, Perception, & Psychophysics*, 77, 1881–1895.
- Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, 1, 261–267. [https://doi.org/10.1016/S1364-6613\(97\)01080-2](https://doi.org/10.1016/S1364-6613(97)01080-2)
- Sperber, R. D., McCauley, C., Ragain, R. D., & Weil, C. M. (1979). Semantic priming effects on picture and word processing. *Memory & Cognition*, 7, 339–345.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522. <https://doi.org/10.1038/381520a0>
- Treisman, A. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12, 242–248.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136. <https://doi.org/10.1016/0010-028590005-5>
- Van der Heijden, A. H. C., La Heij, W., & Boer, J. P. A. (1983). Parallel processing of redundant targets in simple visual search tasks. *Psychological Research*, 45, 235–254.
- VanRullen, R. (2009). Binding hardwired versus on-demand feature conjunctions. *Visual Cognition*, 17(1–2), 103–119.
- VanRullen, R., Reddy, L., & Fei-Fei, L. (2005). Binding is a local problem for natural objects and scenes. *Vision Research*, 45, 3133–3144. <https://doi.org/10.1016/j.visres.2005.05.012>
- Walker, S., Stafford, P., & Davis, G. (2008). Ultra-rapid categorization requires visual attention: Scenes with multiple foreground objects. *Journal of Vision*, 8(4), 21:1–12. <https://doi.org/10.1167/8.4.21>
- Walther, D. B., & Shen, D. (2014). Nonaccidental properties underlie human categorization of complex natural scenes. *Psychological Science*, 25, 851–860. <https://doi.org/10.1177/0956797613512662>
- Wolfe, J. M., & Cave, K. R. (1999). The psychophysical evidence for a binding problem in human vision. *Neuron*, 24, 11–17. <https://doi.org/10.1016/S0896-627380818-1>
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 419–433. <https://doi.org/10.1037/0096-1523.15.3.419>
- Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, 1, 0058. <https://doi.org/10.1038/s41562-017-0058>
- Wyble, B., Folk, C., & Potter, M. C. (2013). Contingent attentional capture by conceptually relevant images. *Journal of Experimental Psychology: Human Perception and Performance*, 39, 861–871. <https://doi.org/10.1037/a0030517>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.