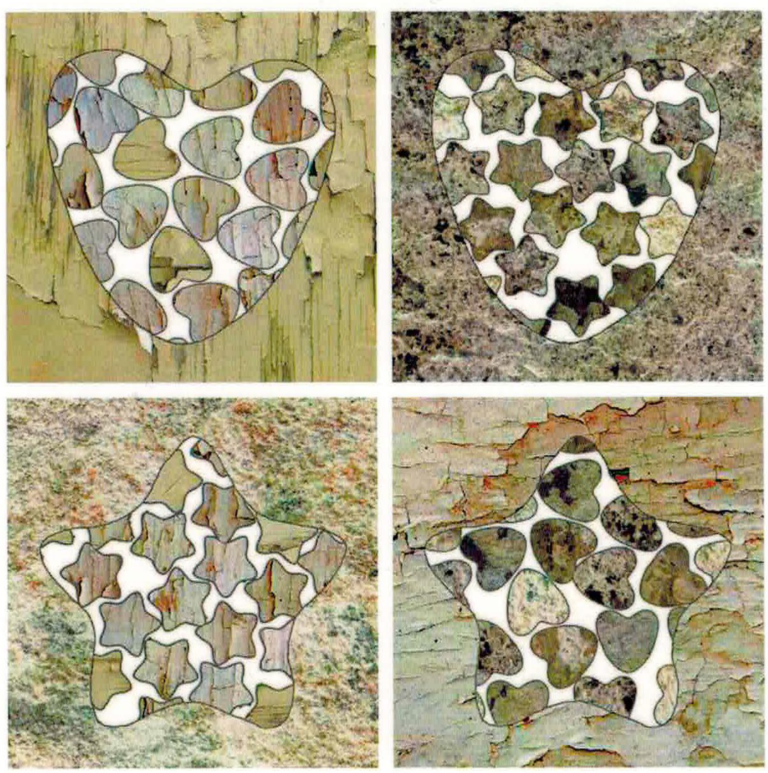
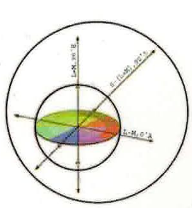
ELSEVIER

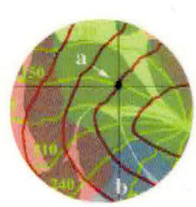
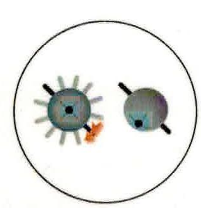
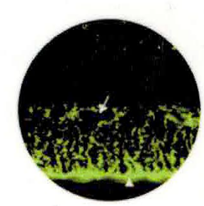
VISION RESEARCH

An International Journal for Functional Aspects of Vision



Biochemistry & Cell Biology • Molecular Biology & Genetics
Anatomy, Physiology, Pathology & Pharmacology • Optics, Accommodation & Refractive Error
Circuitry & Pathways • Psychophysics • Perception • Attention & Cognition
Computational Vision • Eye Movements & Visuomotor Control

ISSN 0042-6989 | Volume 117 | December 2015

Vision Research 117 (2015) 34–40

Contents lists available at ScienceDirect

Vision Research

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

Processing context: Asymmetric interference of visual form and texture in object and scene interactions

Matthew X. Lowe[a,b,*], Susanne Ferber[b,c], Jonathan S. Cant[a]

[a] Department of Psychology at Scarborough, University of Toronto, Toronto, ON M1C 1A4, Canada
[b] Department of Psychology at St. George, University of Toronto, Toronto, ON M5S 3G3, Canada
[c] Rotman Research Institute, Baycrest, Toronto, ON M6A 2E1, Canada

ARTICLE INFO

Article history:
Received 26 February 2015
Received in revised form 30 September 2015
Accepted 5 October 2015

Keywords:
Object–scene perception
Context
Interference
Texture perception
Form perception
Global image features

ABSTRACT

Substantive evidence has demonstrated that scene-centered global image features influence the processing of objects embedded in complex visual scenes. Conversely, a growing body of work suggests that relevant object information may inherently influence diagnostic global scene statistics used in rapid scene categorization. Here, we investigate the potential effects of interference in object–scene perception when attending to form and texture in both simple figure-ground representations and more complex object–background scenes. Results reveal asymmetric interference in the perception of form and texture in object and scene processing: Inconsistent scene texture interfered with the classification of object texture, and inconsistent object form interfered with the classification of scene form, but not vice versa. These findings contribute to our understanding of the interactions between an object and its environment, and further inform our knowledge of the visual features which influence interactivity in object and scene perception.

1. Introduction

A remarkable aspect of the human visual system is the ability to draw on a broad range of cues to rapidly and efficiently identify and categorize objects embedded in a complex visual scene. In general, research has found that knowledge about which objects and scenes tend to co-occur facilitates the efficiency of both the search for and recognition of objects (Biederman, Mezzanotte, & Rabinowitz, 1982; Boyce & Pollatsek, 1992; Davenport, 2007; Davenport & Potter, 2004; De Graef, Christiaens, & d'Ydewalle, 1990; Henderson, Weeks, & Hollingworth, 1999; Joubert, Fize, Rousselet, & Fabre-Thorpe, 2008; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007; Palmer, 1975 for a review, see Oliva and Torralba (2007)). Conversely, a growing body of work has demonstrated evidence for the influence of object information on scene classification through a consistent-object advantage (Davenport, 2007; Davenport & Potter, 2004; Joubert et al., 2007), even without the need to activate semantic information from stored object representations (Mack & Palmeri, 2010). Such research suggests a dual-system, interactive account between scene and object processing. Nevertheless, we currently know very little about the visual features contributing to such an interactive system. Here, we investigate the extent to which common and relatively lower-level visual features (form and texture) influence the interactivity between object and scene processing through visual interference between object and background features.

Scene perception may be governed by general mechanisms that apply broadly across visual processing. For example, seminal work on global processing has suggested that the precedence of global image features is a prevailing property of visual perception, wherein global structure precedes the perception of local elements or fine-grained analyses (Navon, 1977). Navon presented compound letters representing larger figures (global configurations), which were spatially constructed from a suitable arrangement of smaller figures (local elements), and observed an advantage in the processing of global configurations over local elements (i.e., faster responses to global configurations compared with local elements), which he termed the 'global precedence effect'. Critically, when global configurations and local elements were inconsistent, responses to the local elements were subject to interference from the global configurations, but local features did not interfere with global perception. This result was subsequently referred to as the 'global interference effect'. In other words, involuntary attention

* Corresponding author at: Department of Psychology, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON M1C 1A4, Canada.
E-mail address: matthew.lowe@mail.utoronto.ca (M.X. Lowe).

http://dx.doi.org/10.1016/j.visres.2015.10.010
0042-6989/

to the global level was observed when attention was directed to the local level, resulting in global inference in the perception of local elements.

Subsequent research on rapid scene identification has provided support for the primacy of global features over local region and object information, demonstrating that contextual information influencing object–scene interactivity is guided by global image features which direct attention early in the visual processing stream (Torralba, Oliva, Castelhano, & Henderson, 2006). That is not to say, however, that scene and object information is processed in a strictly scene-to-object hierarchical fashion. On the contrary, evidence has demonstrated an advantage for the classification of scenes that contain semantically consistent (compared with inconsistent) objects, suggesting that objects and scenes may be processed interactively and in parallel (Davenport, 2007; Davenport & Potter, 2004; Joubert et al., 2007). In fact, recent research has shown that rapid recognition of a scene's superordinate category of natural or man-made is modulated by the presence of a consistent or inconsistent object, even in the absence of explicit object recognition (Mack & Palmeri, 2010). Mack and Palmeri thus hypothesized that object–scene interference may be more simply explained by relatively low-level visual differences between objects and scenes, as opposed to relatively high-level semantic relationships between them.

Yet which visual features are utilized in such an interactive process? Previous work has shown that coarse, low-level global scene properties determine ultra-rapid scene recognition and categorization (Schyns & Oliva, 1994), and has drawn attention to the roles of form and texture in capturing the diagnostic structure necessary to perform these processes (for a review, see Oliva and Torralba (2006)). Indeed, investigations of scene processing using functional magnetic resonance imaging (fMRI) have demonstrated that the scene-selective parahippocampal place area (PPA), a region shown to respond selectively to scenes over individual objects or faces (Epstein & Kanwisher, 1998), represents scenes by processing global spatial layout (Epstein, Graham, & Downing, 2003). Similarly, recent evidence has revealed that PPA is also sensitive to processing visual cues such as material properties signaled by surface texture (Cant & Goodale, 2007, 2011), suggesting that diagnostic global statistics informing scene identity may incorporate both spatial structure and material properties.

In a similar vein, highly diagnostic visual cues such as surface reflectance properties, surface texture, and surface structure can cue stored knowledge of object material properties such as mass, compliance, and friction (Adelson, 2001; Buckingham, Cant, & Goodale, 2009; Motoyoshi, Nishida, Sharan, & Adelson, 2007). These cues not only aid in visual search and recognition, but also contribute to action planning (Gallivan, Cant, Goodale, & Flanagan, 2014), ultimately affecting how we physically engage with objects of various tactile qualities (e.g., rough vs. smooth), and the adjustment of our gait when moving through an environment containing different surface attributes (e.g., ice vs. grass). Evidence has demonstrated both independence and asymmetric interference in the perception of texture and form in object perception (Cant, Arnott, & Goodale, 2009; Cant, Large, McCall, & Goodale, 2008). In fact, visual texture may be especially important in defining edge and contour information used for finding partially occluded objects in complex and crowded environments (Biederman, 1987). While it has been argued that objects and scenes interact extensively, the influence of visual texture in such an interaction has yet to be explored, despite the importance of texture as a cue in both object and scene processing.

In the present study, we examine the extent to which form and texture consistency influence object–scene interactivity. We focused on global interference effects rather than global precedence effects, since the former capture interactions in visual processing across global and local levels, while the latter simply demonstrates that participants typically process global features faster than local features. In Experiment 1 we aim to initially replicate and extend a global interference effect of form (Navon, 1977) using modified Navon stimuli in simple figure-ground displays, and then investigate this effect for visual texture, predicting a similar interference effect in the perception of texture (i.e., slower judgments of local texture when global and local texture features are inconsistent). Thus, our motivation for Experiment 1 is to validate our stimuli and experimental paradigm by replicating well-established results of global interference in form perception and also to demonstrate novel results of global interference in texture perception. Having done so, we can then extend these findings to the study of more complex object–scene interactions, which we explore in Experiment 2. If form and texture are indeed important visual cues in scene and object perception, and scene perception proceeds from global properties to local elements, we expect to observe global scene interference in the perception of local object properties (i.e., a global interference effect of form and texture). However, as recent evidence has demonstrated that the perception of global scene statistics is modulated by inconsistent object information (Mack & Palmeri, 2010), we will also investigate the potential influence that both object form and object texture have on scene perception. Across both experiments, we elected to focus on speed of processing (reaction time) as a measure of interference, using accuracy only to ensure a constant level of attention across experimental tasks.

2. Experiment 1

Before investigating the interaction between object and scene information in the perception of form and texture, we first aimed to confirm that these visual features are processed in a global-to-local manner. Using modified classic Navon figures (1977), we incorporated both form and texture into simple figure-ground representations (see Fig. 1), predicting a replication of Navon's global interference effect for form (slower judgements of local form when local and global form were inconsistent, but not vice versa), and similarly expected a global interference effect to be found in the perception of texture.

2.1. Participants

Twelve participants (all female) between 20 and 32 years of age ($M = 21.50$) were recruited from the University of Toronto undergraduate community and received course credit for their participation. All participants had normal or corrected-to-normal visual acuity, were right-handed, and gave informed consent in accordance with the University of Toronto Ethics Review Board in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.2. Stimuli and apparatus

Sixty-four stimuli were generated using Adobe Photoshop CS3 software (Adobe Systems, San Jose, CA) and were presented electronically using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) on a ViewSonic 21-in. CRT monitor (1280×1024 resolution; 85-Hz refresh rate). Stimuli subtended $18.4° \times 18.4°$ of visual angle and were presented centrally against a white background following a black central fixation cross (subtending $1° \times 1°$) at a viewing distance of 52 cm. The stimuli were constructed so that visual features (form: heart versus star; texture: paint versus rock) could vary at both global and local levels of attention, and importantly, variations in each feature were manipulated across

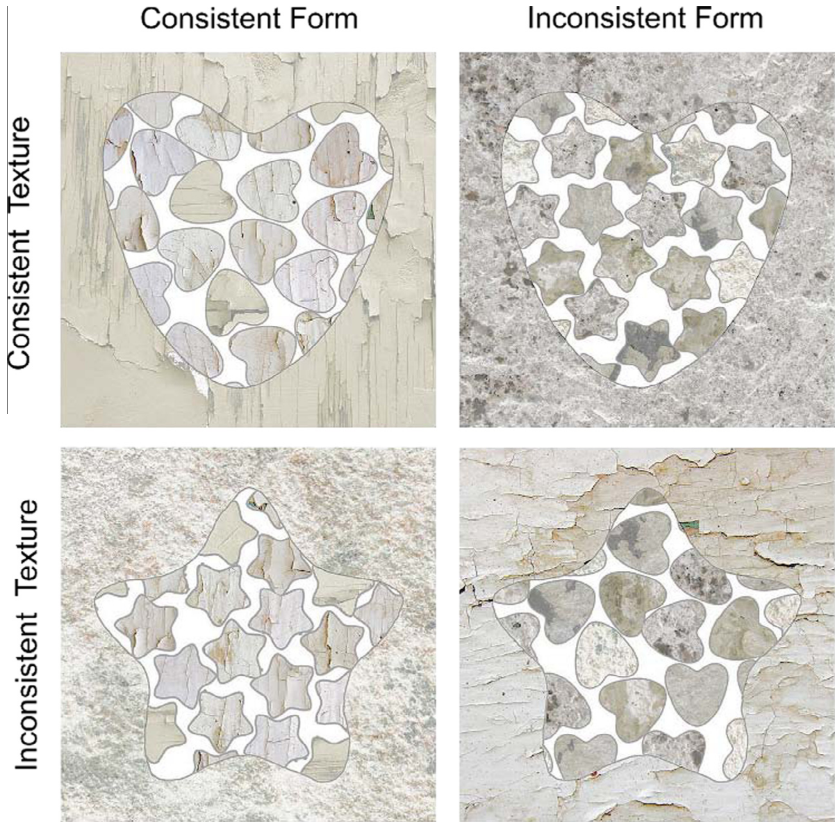


Fig. 1. Examples of the stimuli used in Experiment 1. The stimuli could vary along two features (form and texture), two levels of consistency (consistent and inconsistent) and two scopes of attention (global and local); global form: large shape; local form: small shapes contained within the large shape; global texture: textured surface behind the global shape; local texture: textured surfaces on the local shapes.

levels (consistent: similar global and local shape, similar global and local texture; inconsistent: different global and local shape, different global and local texture). Variations in each visual feature were matched at both global and local levels (i.e., 32 instances of global star, 32 instances of global heart, 32 instances of global paint, 32 instances of global rock, etc.). In order to avoid responses based on any one local element, local texture elements from the same category (i.e., paint or rock) were heterogeneous in nature and the locations and orientations of both local form and texture elements were jittered across stimuli. To ensure that observer classification across texture categories was independent of visually distinctive color cues, the chromaticity of each stimulus' texture was calculated using Matlab software (MathWorks, Natick, MA) and modified through level adjustment using Adobe Photoshop CS3 (see Supplementary Materials and Supplementary Fig. 1 for more details). Mean luminance for all stimuli averaged 202.10 (on a 0–255 luminance scale) with a standard deviation of 9.49.

2.3. Design and procedure

The experiment contained four blocks of trials representing conditions of form and texture perception at both a global and local scope of attention (i.e., global form; global texture; local form; local texture). Before beginning the experiment, participants were given five practice trials per condition to become familiar with the task. Each block began with an instruction to attend to either "shape" or "texture" at either global- or local-levels of attention (instructed as either "large" or "small," respectively). Following an initial key press, each trial began with a central fixation cross displayed for 2000 ms, after which the stimulus was presented and remained on screen until response. Participants were instructed to make a speeded classification (shape: heart or star; texture: paint or rock) after the onset of the stimulus using either the "1" or "2" keys on the number pad of a computer keyboard, which would then terminate the trial. For the experiment proper, each block contained sixty-four randomly presented stimuli with equal numbers of consistent and inconsistent trials. The order of presentation of the four blocks was counterbalanced across participants. Each block was separated by an instruction screen informing participants that they may take a short break, and reminding them to respond as quickly and accurately as possible in the following block. Accuracy and response latency was recorded for each trial.

2.4. Results and discussion

Participants made few errors overall, and accuracy was consistently high across conditions (see Table 1). An initial outlier analysis was performed separately on each participant, and response latencies 2.5 standard deviations above or below the mean reaction time for each condition were excluded from the analysis. Response latencies were analyzed using a three-way repeated measures analysis of variance, with scope of attention (global vs. local),

Table 1
Average accuracy (percent correct) in each condition for both experiments.

Condition	Experiment 1	Experiment 2
Global form	96.48	96.93
Local form	95.44	97.33
Global texture	94.01	96.04
Local texture	93.75	96.04

feature (form vs. texture), and consistency (consistent vs. inconsistent) as factors. Pairwise post-hoc comparisons were performed using the Bonferroni–Holm procedure to correct for inflations of the Type I error rate due to multiple comparisons (alpha = 0.05). This analysis procedure was performed in both experiments.

Mean response latencies can be seen in Fig. 2 as a function of scope of attention, feature, and consistency. Significant main effects were found for feature [form $M = 541.4$ ms; texture $M = 571.6$ ms; $F(1,11) = 12.40$, $p = .005$] and consistency [consistent $M = 549.0$ ms; inconsistent $M = 564.0$ ms; $F(1,11) = 14.65$, $p = .003$], and a significant attention-by-consistency interaction [$F(1,11) = 5.06$, $p = .046$] was also found. Further planned pairwise comparisons between consistent and inconsistent trials for each of the four conditions of interest (i.e., global form, local form, global texture, and local texture) were performed using the Bonferroni–Holm correction for multiple comparisons.

The results of Experiment 1 provide clear evidence for the primacy of global percepts of both form and texture as measured through global interference: we found slower response latencies for classifications of local features when local and global features were inconsistent [local form: consistent $M = 527.6$ ms; inconsistent $M = 557.4$ ms; $t(11) = 3.21$, $p = .033$; local texture: consistent $M = 565.1$ ms; inconsistent $M = 585.9$ ms; $t(11) = 2.92$, $p = .042$], but not vice versa [global form: consistent $M = 534.9$ ms; inconsistent $M = 545.8$ ms; $t(11) = 1.95$, $p = .16$; global texture: consistent $M = 568.5$ ms; inconsistent $M = 567.1$ ms; $t(11) = .80$, $p = .44$]. The main effect of scope of attention was not significant [$F(1,11) = .55$, $p = .47$], and no difference in response latency between global and local attention was found in both form and texture processing [global form $M = 540.3$ ms; local form $M = 542.5$ ms; $t(11) = .77$, $p = .46$; global texture $M = 567.8$ ms; local texture $M = 575.5$ ms; $t(11) = .68$, $p = .51$], yet a global interference effect was observed for both features. That is to say, global interference for both form and texture was observed despite no significant overall differences in response latencies between global and local attention conditions for both features (i.e., no global precedence effects). This indicates that interference from the global percept is not dependent on slower response latencies for local compared with global elements.

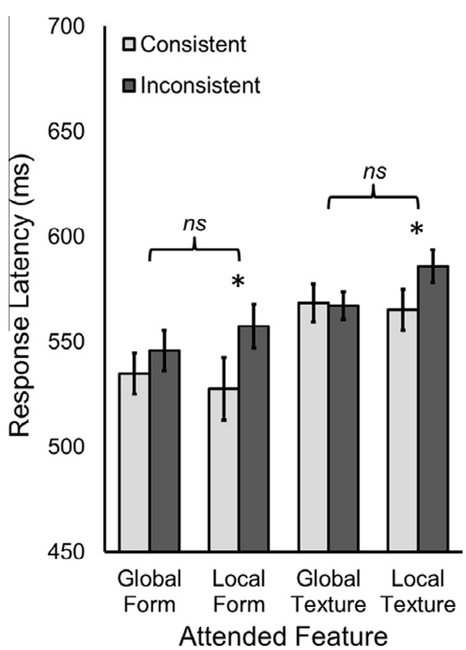


Fig. 2. Results of Experiment 1 for each condition (global form; local form; global texture; local texture). Light bars represent consistent global and local features, and dark bars represent inconsistent global and local features. All statistical comparisons are between the consistent and inconsistent conditions of each attended feature, except those denoted by the parenthesis, which compared global versus local processing within each feature (i.e., form and texture), collapsed across levels of consistency. Results are based on data from 12 participants, in a repeated-measures design. Error bars represent the standard error of the mean. $^*p < 0.05$.

Finally, when the consistency of the unattended feature was held constant, significant effects were maintained across both local form and local texture conditions [$t(11) = 2.87$, $p = .030$, and $t(11) = 2.78$, $p = .018$, respectively], indicating that differences in mean latency between consistent and inconsistent trials were driven by changes of consistency in the attended feature, independent of changes in the unattended feature. These results contribute a novel and important finding regarding the primacy of global over local surface-texture features in figure-ground perception, independent of form. This last finding is consistent with the independence of form and texture in the perception of single objects (Cant et al., 2008).

3. Experiment 2

In the previous experiment, we replicated Navon's (1977) findings of a global interference effect in form perception, revealing significant effects of consistency when attending to local but not global form. Moreover, the results of Experiment 1 introduced the novel finding of a global interference effect in texture perception, revealing that global texture properties precede and subsequently influence the processing of local texture properties. In the present experiment, we examined whether these findings would generalize to more complex object–scene environments. To investigate this, a single object was placed centrally within a scene upon a raised platform with a neutral background (see Fig. 3). We incorporated a neutral background to ensure that participants could clearly perceive both scene and object features. Indeed, the lack of a neutral background on the raised platform would cause the visual features of the object and scene to blend together, and would thus make perception and classification of the object features much more difficult. Using this design thus allowed us to examine interference effects between object and scene processing while avoiding potential object individuation difficulties within a scene.

3.1. Methods

3.1.1. Participants

Ten new participants (six females) between 18 and 21 years of age ($M = 19.80$) were recruited from the University of Toronto undergraduate community and received course credit for their participation. All participants had normal or corrected-to-normal visual acuity, were right-handed, and gave informed consent in accordance with the University of Toronto Ethics Review Board in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

3.1.2. Apparatus, stimuli and procedure

The apparatus used in this experiment was identical to the apparatus used in the previous experiment, with the addition of Blender 2.0 software (Stichting Blender Foundation, Amsterdam), which was used to render 3-dimensional indoor environments and generate stimuli. One-hundred and twenty-eight new stimuli, each subtending $33.45° \times 21.28°$ in visual angle, were created. During the experiment each of these new stimuli was presented centrally against a white background following a black central fixation cross (subtending $1° \times 1°$) at a 52 cm viewing distance. To maintain consistency with the previous experiments, object texture was counterbalanced to contain equal representations of homogenous stimuli (object–scene textures were selected from

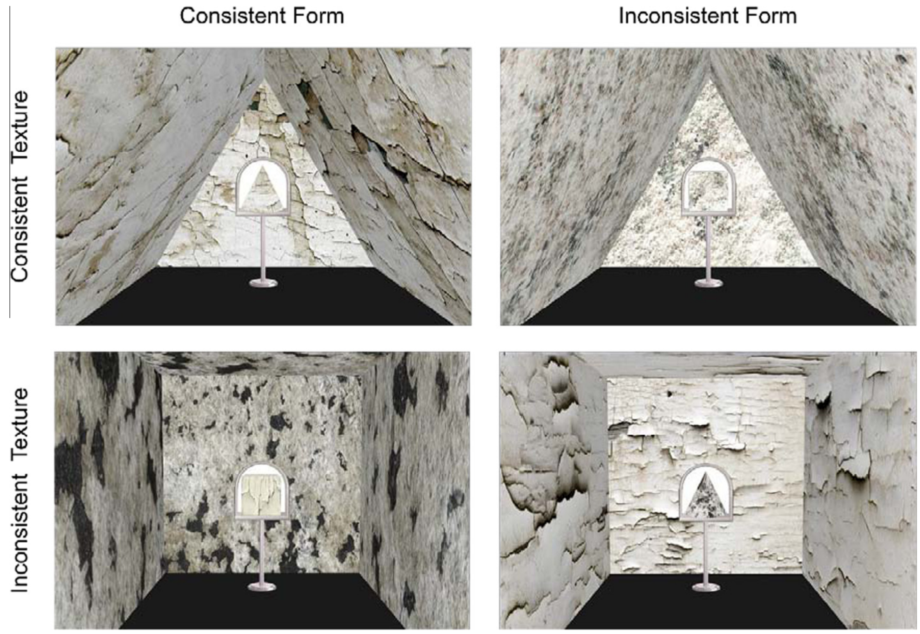


Fig. 3. Examples of the stimuli used in Experiment 2. Stimuli could vary along two features (form and texture), two levels of consistency (consistent and inconsistent) and two scopes of attention (scene and object).

the same source image) and heterogeneous stimuli (object–scene textures were selected from different source images from the same texture category; i.e., paint or rock). Images were rendered using a constant view-point with consistent lighting across feature conditions in order to maintain overall consistency in surface area, perspective, and reflectance. The stimuli were created to have variations in visual features (form: square versus triangle; texture: paint versus rock) at both a global (scene) and local (object) processing level, and contained either consistent or inconsistent features across levels. Variations in each visual feature were matched at both global and local levels (i.e., 64 instances of global square, 64 instances of global triangle, 64 instances of global paint, 64 instances of global rock, etc.). Mean luminance for all stimuli averaged 140.49 (on a 0–255 luminance scale) with a standard deviation of 28.57. The experimental design and procedure were identical to those used in the previous experiment, except for the fact that more trials were included in this experiment, and participants were now instructed to attend to either 'scene shape', 'scene texture', 'object shape', or 'object texture'. To avoid any ambiguity, scene features were explicitly defined as those belonging to the walls of the room, and object features were defined as those belonging to the object sitting on the pedestal near the back wall of the room.

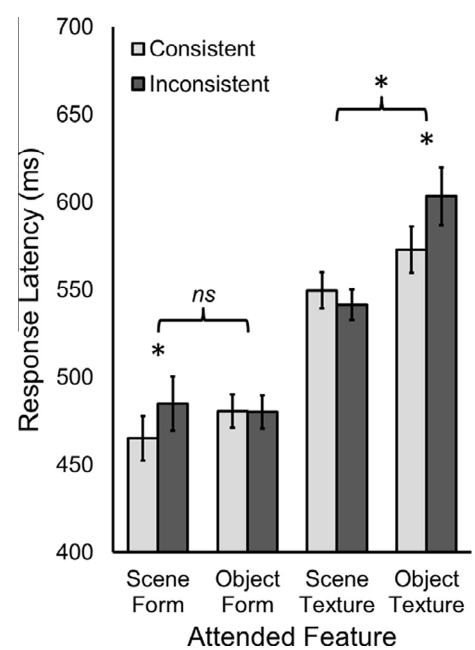


Fig. 4. Results of Experiment 2 for each condition (scene form; object form; scene texture; object texture). Light bars represent consistent scene and object features, and dark bars represent inconsistent scene and object features. All statistical comparisons are between the consistent and inconsistent conditions of each attended feature, except those denoted by the parenthesis, which compared scene versus object processing within each feature (i.e., form and texture), collapsed across levels of consistency. Results are based on data from 10 participants, in a repeated-measures design. Error bars represent the standard error of the mean. $^{*}p < 0.05$.

3.2. Results and discussion

In order to eliminate any potential difference in response latencies driven by differences in the luminance of the images across conditions, twenty-seven stimuli were removed prior to data analysis following a luminance outlier analysis (see Supplementary Material and Supplementary Fig. 2 for details). Mean luminance for all remaining stimuli averaged 147.90 (on a 0–255 luminance scale) with a standard deviation of 20.85.

As we observed in Experiment 1, participants made very few errors overall (see Table 1). Mean response latencies can be seen in Fig. 4 as a function of scope of attention, feature, and consistency. A significant main effect of feature was found [form $M = 477.7$ ms; texture $M = 566.7$ ms; $F(1,9) = 24.47$, $p = .001$], as well as a significant three-way interaction between scope of attention, feature, and consistency [$F(1,9) = 19.57$, $p = .002$]. Similar to Experiment 1, planned pairwise comparisons for each of the four conditions of interest were performed using the Bonferroni–Holm correction for multiple comparisons.

Results of the present experiment demonstrate significantly lower mean response latencies for judgements of object texture when the background scene texture was consistent, compared with when it was inconsistent [consistent $M = 572.8$ ms; inconsistent $M = 603.3$ ms; $t(9) = 3.69$, $p = .020$], but not vice versa [judgments of scene texture: consistent $M = 549.4$ ms; inconsistent $M = 541.3$ ms; $t(9) = 1.08$, $p = .614$]. In addition, we observed longer response latencies when attending to object texture compared to scene texture [global texture $M = 545.4$ ms; local texture $M = 588.1$ ms; $t(9) = 2.72$, $p = .024$], consistent with Navon's (1977) global precedence hypothesis. These results replicate and extend the results of a global interference effect in texture perception observed in Experiment 1, and indicate that global scene texture may form a contextual cue in influencing object perception and recognition through the primacy of global scene statistics.

Although no difference in response latencies was observed between scene and object form [global form $M = 475.0$ ms; local form $M = 480.4$ ms; $t(9) = .97$, $p = .36$], we observed interference of object form on classifications of scene form [consistent $M = 465.2$ ms; inconsistent $M = 484.9$ ms; $t(9) = 3.16$, $p = .034$], but not vice versa [judgments of object form: consistent $M = 480.6$ ms; inconsistent $M = 480.2$ ms; $t(9) = .05$, $p = .958$] (i.e., a local, or object interference effect). These findings are consistent with previous evidence demonstrating a deleterious effect of object information on scene identification when object information is inconsistent with the scene (Davenport & Potter, 2004; Joubert et al., 2007; Mack & Palmeri, 2010). Interestingly, this effect was not found for scene texture classification, indicating both independence and asymmetric feature-specific interference in form and texture processing. These results are largely consistent with previous investigations demonstrating independent processing and asymmetric interference of form and texture in object perception (Cant et al., 2008, 2009), yet extend these findings to object–scene interactions. Similar to the results in Experiment 1, when the consistency of the unattended feature was held constant, the significant effects of both local texture [$t(9) = 2.90$, $p = .035$] and global form [$t(9) = 2.35$, $p = .043$] were maintained.

4. General discussion

The present study investigated the potential effects of visual form and texture interference in object and scene perception. Using simple figure-ground images (i.e., Experiment 1), we confirmed a global interference effect in the perception of local form, and subsequently extended these findings to the perception of texture. Examination of more complex scenes (i.e., Experiment 2), however, revealed asymmetric interference between these visual features. Specifically, our findings demonstrated that inconsistent scene texture interfered with the classification of object texture, and inconsistent object form interfered with the classification of scene form, but not vice versa. Taken together, these data suggest that relatively lower-level visual differences may reveal patterns of asymmetric interference in object and scene processing, and consequently influence object–scene interactivity.

4.1. Asymmetric interference of visual form and texture

Models accounting for rapid scene categorization have drawn attention to the contributions of both form and texture in global scene statistics (for a review, see Oliva and Torralba (2006)). Interestingly, it has been shown that an individual with profound form vision deficits was able to accurately classify scenes using texture and color information alone (Steeves et al., 2004), suggesting an asymmetry in how form and texture are utilized in natural scene perception. In the present study, the asymmetry observed between these features further suggests fundamental differences in how form and texture interact in object–scene processing. Although our results were consistent with previous research demonstrating faster classification of form relative to texture (e.g. Cant et al. (2008)), differential speeds of processing are an unlikely explanation for the observed asymmetry. For instance, this argument does not account for the observed global interference of texture, but not form, in scene perception (i.e., Experiment 2), nor would it account for the utilization of both these features in ultra-rapid scene identification (Schyns & Oliva, 1994). Here, we propose an alternate account for this asymmetry through a model of scene categorization based on the interaction of low-level object–scene features (Mack & Palmeri, 2010).

To account for a consistent-object advantage in scene classification (Davenport, 2007; Davenport & Potter, 2004; Joubert et al., 2007, 2008), Mack and Palmeri (2010) developed a model based on the assumption that global-based representations are influenced by objects containing visual properties which shift the global representation of a scene away from the expected visual regularities. With this in mind, the perception of global scene structure may be perceived as a coherent and holistic global feature within which object properties (i.e., form) are integrated. Inconsistent object information from the environment may therefore interfere with scene categorization by influencing global image statistics which contribute to the perception of scene structure. This local interference as a result of object properties would further explain why object, but not scene, interference was observed in the perception of visual form. Indeed, object form information may be processed relatively automatically and definitively without requiring the need for global scene input, whereas scene form may build on the presence of local components to derive a global representation of the visual environment.

In sharp contrast, due to the relative uniformity and redundancy of elements which constitute visual texture, local object outliers may be suppressed, resulting in a bias in favor of the global average scene texture representation consistent with the well-established global precedence hypothesis (Navon, 1977). In other words, a bias toward visually dominant global texture statistics, and away from local outliers, would explain why we did not observe a local interference effect for texture in scene perception (i.e., Experiment 2). Interestingly, such reduced weighting of outliers when computing the perceived average has been demonstrated in the representation of ensemble statistics (Haberman & Whitney, 2010; for a review, see Alvarez (2011)), which, just like the representation of texture, have been shown to rely on neural processing in scene-selective visual cortex (Cant & Xu, 2012). Thus, a close-knit relationship may exist between these two types of visual processes and scene representation, all of which may be mediated by similar cognitive and neural mechanisms. Taken together, these findings suggest that, although form and texture appear globally-dominant in simpler stimulus configurations (i.e., Experiment 1), exploring their processing in more naturalistic settings (i.e., Experiment 2) reveals that the importance of local features may differ in object–scene interactions through their relative contributions to global scene statistics, likely via mechanisms which relate to the feature-specific integration of local components with global statistics. Specifically, where local object structural information may inform global scene structure, global and visually dominant scene texture may minimize local object texture interference.

4.2. Texture in the context of a scene

Our findings have shown that surface properties such as texture can form a contextual link between the processing of objects and scenes, and this interactive processing proceeds from the scene-centered to object-centered scale of attention. Yet how important is this interaction? Texture is instrumental in providing the visual cues necessary to infer the material properties of an object (natural vs. manufactured; heavy vs. light), which subsequently aids in identification and action planning necessary for interacting with objects in our environment (Buckingham et al., 2009; Gallivan et al., 2014). Indeed, the search for and recognition of objects in the natural world often requires knowledge about relevant material properties, derived from surface-based cues such as texture, especially when form is degraded through occlusion or is uninformative (Biederman & Ju, 1988). These cues may therefore be highly influential in drawing attention to contextually-relevant objects and establishing a relationship between an object and its environment. In light of the global interference effect for texture in object–scene stimuli observed in Experiment 2, it appears that contextual information extracted from environmental texture cues can facilitate such search and recognition through knowledge about real-world scene categories (e.g., natural vs. man-made), especially when the semantic properties of an object are not easily accessed. Interestingly, in addition to texture (e.g. Cant and Goodale (2007)) and scene perception (e.g. Epstein and Kanwisher (1998)), parahippocampal cortex has been implicated in processing contextual associations (e.g. Bar, Aminoff, and Schacter (2008)). Thus, the present results provide a bridge between the processing of lower-level visual features and higher-level contextual associations that work together to influence object and scene perception. While these findings provide a promising framework for future research investigating the interactions of common visual features in object–scene processing, there is some degree of speculation surrounding the mechanisms underlying these interactions, and the neural processes at play. The present results thus provide a unique opportunity for future research to investigate the novel hypotheses generated here in order to more fully understand the processes surrounding visual feature asymmetry and interference in object–scene dynamics.

5. Conclusion

Our main findings have revealed asymmetric interference in the perception of form and texture in object and scene processing. Specifically, inconsistent object form interfered with scene form classification, while inconsistent scene texture interfered with object texture classification, but not vice versa. These findings extend an existing body of work on contextual associations between object and scene processing, and further highlight the role of lower-level visual features in object–scene interactivity. The data presented here therefore pave the way for future avenues of research investigating the relative weighting of various visual features on the processing of objects within scenes.

Acknowledgments

This research was supported by a Natural Sciences and Engineering Research Council (NSERC) Discovery Grant awarded to J.S.C., and Grants from the Canadian Institute of Health Research (CIHR) and NSERC awarded to S.F.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.visres.2015.10.010.

References

Adelson, E. H. (2001). On seeing stuff: The perception of materials by humans and machines. In *Photonics West 2001-electronic imaging* (pp. 1–12). International Society for Optics and Photonics.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in cognitive sciences, 15*(3), 122–131.

Bar, M., Aminoff, E., & Schacter, D. L. (2008). Scenes unseen: The parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *The Journal of Neuroscience, 28*, 8539–8544.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115.

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology, 20*(1), 38–64.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*(2), 143–177.

Boyce, S. J., & Pollatsek, A. (1992). Identification of objects in scenes: The role of scene background in object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(3), 531.

Buckingham, G., Cant, J. S., & Goodale, M. A. (2009). Living in a material world: How visual cues to material properties affect the way that we lift objects and perceive their weight. *Journal of Neurophysiology, 102*(6), 3111–3118.

Cant, J. S., Arnott, S. R., & Goodale, M. A. (2009). FMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. *Experimental Brain Research, 192*(3), 391–405.

Cant, J. S., & Goodale, M. A. (2007). Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cerebral Cortex, 17*(3), 713–731.

Cant, J. S., & Goodale, M. A. (2011). Scratching beneath the surface: New insights into the functional properties of the lateral occipital area and parahippocampal place area. *The Journal of Neuroscience, 31*(22), 8248–8258.

Cant, J. S., Large, M. E., McCall, L., & Goodale, M. A. (2008). Independent processing of form, colour, and texture in object perception. *Perception, 37*, 57–78.

Cant, J. S., & Xu, Y. (2012). Object ensemble processing in human anterior-medial ventral visual cortex. *The Journal of Neuroscience, 32*(22), 7685–7700.

Davenport, J. L. (2007). Consistency effects between objects in scenes. *Memory & Cognition, 35*(3), 393–401.

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science, 15*(8), 559–564.

De Graef, P., Christiaens, D., & d'Ydewalle, D. (1990). Perceptual effects of scene context on object identification. *Psychological Research, 52*, 317–329.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392*(6676), 598–601.

Epstein, R., Graham, K. S., & Downing, P. E. (2003). Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron, 37*(5), 865–876.

Gallivan, J. P., Cant, J. S., Goodale, M. A., & Flanagan, J. R. (2014). Representation of object weight in human ventral visual cortex. *Current Biology, 24*(16), 1866–1873.

Haberman, J., & Whitney, D. (2010). The visual system discounts emotional deviants when extracting average expression. *Attention, Perception, & Psychophysics, 72*(7), 1825–1838.

Henderson, J. M., Weeks, P. A., Jr, & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 210.

Joubert, O. R., Fize, D., Rousselet, G. A., & Fabre-Thorpe, M. (2008). Early interference of context congruence on object processing in rapid visual categorization of natural scenes. *Journal of Vision, 8*(13), 11.

Joubert, O. R., Rousselet, G. A., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research, 47*(26), 3286–3297.

Mack, M. L., & Palmeri, T. J. (2010). Modeling categorization of scenes containing consistent versus inconsistent objects. *Journal of Vision, 10*(3), 11.

Motoyoshi, I., Nishida, S. Y., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature, 447*(7141), 206–209.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*(3), 353–383.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research, 155*, 23–36.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11*(12), 520–527.

Palmer, T. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3*, 519–526.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science, 5*(4), 195–200.

Steeves, J. K., Humphrey, G. K., Culham, J. C., Menon, R. S., Milner, A. D., & Goodale, M. A. (2004). Behavioral and neuroimaging evidence for a contribution of color and texture information to scene classification in a patient with visual form agnosia. *Journal of Cognitive Neuroscience, 16*(6), 955–965.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*(4), 766.