Natural scenes upset the visual applecart

Jochen Braun

Plymouth Institute of Neuroscience, Devon PL4 8AA, UK

The effortless ease of everyday vision seems to contradict numerous findings on the limited capacity of visual attention. However, natural scenes appear to escape the stringent limitations of attention that apply to seemingly far simpler stimuli. This astonishing result will oblige us to rethink the nature of visual attention and its limited capacity.

Everyday vision differs in many ways from what is commonly studied in vision laboratories. For one, the complex natural scenes that inundate everyday vision are far cry from the meagre fare of geometric shapes (letters, gratings, etc.) served in most laboratories. For another, the stimuli of everyday vision are rarely as repetitious as those used in the laboratory, raising the spectre of narrow skills more highly practised than their everyday counterparts. The list of differences lengthens further when we consider the dynamic aspects of stimulation, visuomotor integration, and so forth. These are sobering thoughts for anyone hoping to understand vision in the real world, for we cannot expect that laboratory findings will necessarily generalize to the domain of ultimate interest.

A particularly dramatic failure of vision to generalize from geometric shapes to natural scenes has now been reported by Li, Van Rullen, and colleagues [1]. Their key finding is that observers categorize the gist of natural scenes (e.g. whether they contain an animal, a vehicle, or neither) even though they cannot, under otherwise identical conditions, distinguish simple geometric shapes. Thus it would appear that natural scenes are processed more efficiently, and with less call on attentional resources, than some simple geometric shapes. Before accepting this startling conclusion, one will wish to take a closer look at the details. Li and colleagues flash natural scenes briefly in the visual periphery, followed by a custom-made stochastic mask to limit visual persistence. The effective presentation times (≤ 80 ms) fall in the ‘ultra-rapid categorization’ range described by Thorpe and colleagues, among others [2–4]. With full attention, observers categorize approximately 3 out of 4 of scenes correctly, that is to say, they correctly distinguish scenes with and without animals or scenes with and without vehicles in about 75% of all trials (Fig. 1a). The same level of performance is obtained for discriminating the simple geometric shapes that serve as a comparison (rotated Ts or Ls, two-coloured disks, Fig. 1b,c), although presentation times need to be longer, hinting that

![Fig. 1](https://tics.trends.com)

---


PII: S1364-6613(02)00012-8

Corresponding author: Jochen Braun (achim@pion.ac.uk).

http://tics.trends.com
these ‘simple’ discriminations may not be so simple after all.

**Superiority of natural scenes**

A yawning gulf separates natural images and geometric shapes, however, when the observer’s attention is coaxed away from either of these stimuli. For natural images continue to be categorized with the same success rate even in the ‘near absence’ of attention, whereas the discrimination of geometric shapes collapses to chance level. This unexpected robustness of natural image processing is especially impressive in view of the demonstrable effectiveness of the attentional coaxing [5]. Specifically, observers focussed attention on a cluster of randomly positioned and oriented letters near fixation and performed a visual search task with respect to these stimuli. Thus, observers reported independently on two separate regions of the display (central search and peripheral categorization/discrimination), but focussed attention as much as possible on the central region alone. To verify that observers maintained this lopsided distribution of attention, central performance was compared against the baseline level of performing the central task alone (i.e. ignoring the peripheral tasks) and no significant difference was found. Any reduction of attention would, of course, have been reflected in a loss of central performance. Another important aspect of the manipulation was to engage attention for an extended interval, which generally started well before and ended well after the peripheral stimulation. Accordingly, attention could not have ‘swerved’ to the periphery either before or after focussing on the central array. In short, the authors provide solid buttressing for their claim to ‘near absence’ of attention, at least on the conventional definition of attention as a limited capacity (see below).

Leaving the implications for attention aside for the moment, one is puzzled why natural images should be so much more readily distinguishable than geometric shapes in the situation at issue. Indeed, this is not the first indication that natural images may trigger a more profound perceptual and neural response than plain geometric shapes. Mack and Rock studied the extent to which an observer is aware of highly visible but entirely unexpected stimuli while his or her attention focuses elsewhere in the display [6], a rather similar situation to the one under discussion. The results show a dramatic failure of awareness when the unexpected stimulus is a geometric shape (25–80% ‘blindness’ for shapes of various sorts), but virtually complete awareness of natural scenes (including faces). In fact, none of the observers in this study failed to notice the unexpected appearance of a natural scene, and most gave fairly accurate descriptions of its contents, in stark (and unexplained!) contrast to geometric shapes. Accordingly, it may well be generally true that natural images remain highly discriminable under conditions of diminished attention. The underlying reason may simply be the nature of the neural code for visual information. In single-unit recordings in area V1 of macaque, natural stimulation is found to increase both the sparseness and the intensity of V1 responses, coupling intense bursts with lower spiking activity overall and thus increasing information per spike and efficiency per neuron [7,8]. Although methodological differences prevent a definitive conclusion at this time, it would appear that information transmission by V1 neurons (median values of bits/sec and bits/spike) varies considerably with stimulus type and may indeed be substantially higher for natural scenes than for geometric stimuli such as gratings [8,9]. Theoretical research has shown that the informative components of natural scenes are sparsely distributed in space and time [10] and it is widely thought that the neural code takes advantage of this fact. If so, the neural representation of natural scenes will be more intense and more intermittent than that of other stimuli that appear ‘simple’ to us. In short, natural scenes might be intrinsically superior stimuli to the ‘simple’ geometric shapes commonly used in vision research.

**Vision without attention**

In addition to the possible special status of natural images, the present results also bear on our visual capabilities in the ‘near absence’ of attention. It has long been known that trained psychophysical observers are aware of salient stimuli ‘outside the focus’ of attention and are able to discriminate some of their attributes (e.g. contrast, colour, orientation, etc.) although more complex attributes remain indistinguishable (e.g. identity of rotated Ts and Ls, colour arrangement of two-coloured disks) [11,12]. More recently, the prior familiarity of observers with flashed, masked display was found to be important to these findings [13,14]. In particular, vision ‘outside the focus’ of attention was confirmed for ‘expert’ but not for ‘novice’ observers (the specific displays in question were novel to both groups, but the ‘experts’ had some prior experience with flashed, masked displays).

Importantly, Li et al. demonstrate that their findings are not due to previous familiarity or training. Firstly, scene categorization in the ‘near absence’ of attention occurs also for unfamiliar scenes, which observers have not seen during their extensive training. Secondly, observers can alternate between categorizing animals and categorizing vehicles without adverse effects on performance, suggesting that categorization extends even to unfamiliar categories (Fig. 1a). Finally, Li et al. show that the discrepancy between natural scenes and geometric shapes persists even with equal training on each stimulus type.

In perceptual research, ‘absence’ of attention remains a controversial notion with no operational definition that is universally agreed. One school of thought relies on the performance effects of dividing attention (not necessarily equally) between two conflicting tasks [15,16]. As one task receives more attention, the other task necessarily receives less, causing the respective performance levels to move in opposite directions. When one task is performed at chance level, we can infer ‘absence’ of attention
have known that observers process natural scenes
by paying attention to their 'focus of attention'. For
about the same amount of time, we have known
that natural stimuli ‘capture’ attention more
efficiently than geometric shapes. However, the
search was performed equally well with both
peripheral stimuli, providing no evidence that
attention capture ‘detracted’ differentially from the
central search. Accordingly, a ‘captured attention’
account would not be consistent with a limited
capacity of visual attention [15,16].

Neural theory of attention
From a neural point of view, the notion of ‘unattended’
stimuli is less problematic. The dominant neural
theory of attention is Biased Competition Theory
(BCT), which views attention as a dynamic state
of activity in a hierarchy of visual cortical (and possibly
sub-cortical) areas [19,20]. In BCT, the activity pattern
is shaped by visual input signals, by competitive
interactions within the hierarchy, and by ‘bias signals’
from higher areas. The pattern that emerges might
comprise multiple winners (‘salient objects’), all of
which become available to awareness, short-term
memory, voluntary reporting, and so on. Objects that
gained saliency from ‘bias signals’ would be considered
‘attended’ in the perceptual terminology, whereas
objects that gained saliency exclusively through bot-
tom-up interactions would be considered ‘unattended’.
It follows that conscious vision ‘outside the focus of
attention’ should embrace only the most competitive
stimuli in the field of view. In the ‘near absence’ of
attention created by Li et al., observers would have
been aware of the single peripheral stimulus due to its
competitive nature. It might be worth pointing out that
this awareness should not extend to multiple periph-
ernal stimuli, where competitive advantage and con-
scious access should be lost. In other words, it should
not be possible to search an entire array of natural
scenes without focussed attention. Rather, visual
search with natural scenes should still be ‘serial’ and
search times should increase with array size.

Conclusion
For more than 20 years, the effortless ease of everyday
vision has perched uneasily next to some cherished
laboratory findings on visual capacity limitations (e.g.
the ponderous pace of serial visual search). To resolve this
paradox, it has even been suggested that the experience of
a rich visual world in front of our eyes is a ‘grand illusion’
sustained only by memory. On a less grand but still helpful
scale, it has been shown that visual capacity is more
commodious than at first apparent, thanks to bottom-up
interactions contributing visual capabilities ‘outside the
focus’ of attention. For about the same amount of time, we
have known that observers process natural scenes
extremely rapidly, at least to the point of recognising
‘gist’ [2–4].

Li et al. have now gone even further by demonstrating
that natural scenes escape some of the stringent capacity
limitations that pertain to geometric shapes (i.e. the
stimuli used most commonly in visual psychophysics). If
visual capacity is even more commodious for natural
scenes, this may reconcile the results of visual psycho-
physics with our experience of everyday vision. Of course,
this resolution would also oblige us to revisit many
results on capacity limitations previously obtained with
geometric shapes. In short, everyone working on visual
attention and capacity limitations ought to be grateful to
Li, Van Rullen and colleagues for having well and truly
upset the apple cart.

References
2 Fabre-Thorpe, M. et al. (2001) A limit to the speed of processing in
13, 171–180
sequence of pictures. J. Exp. Psychol. 81, 10–15
perception: detecting and judging objects undergoing relational
violations. Cogn. Psychol. 14, 143–171
undifferentiated: concurrent discrimination of form, color, and motion.
Percept. Psychophys. 61, 1241–1255
7 Vinje, W.E. and Gallant, J.L. (2000) Sparse coding and decorrela-
tion in primary visual cortex during natural vision. Science 287,
1273–1276
nonclassical receptive field increases information efficiency in V1.
J. Neurosci. 22, 2904–2915
9 Reich, D.S. Mechler, F. and Victor, J.D. (2000) Formal and attribute-
specific information in primary visual cortex. J. Neurophysiol. 57,
132–146
statistics and neural representation. Annu. Rev. Neurosci. 24,
1193–1216
Percept. Psychophys. 48, 45–48
12 Braun, J. and Julesz, B. (1998) Withdrawing attention at little or no
cost: detection and discrimination tasks. Percept. Psychophys. 60,
1–23
requirements in a ‘preattentive’ feature search task. Nature 387,
805–807
424–425
15 Sperling, G. and Melchner, M.J. (1978) The attention operating
characteristic: examples from visual search. Science 202, 315–318
dynamics of spatial attention. Psychol. Rev. 102, 503–532
Trends Cogn. Sci. 4, 147–155
245–277
mechanisms subserve attention in macaque area V2 and V4.
J. Neurosci. 19, 1736–1753
attention: integration of experimental single-cell and fMRI data.
Cereb. Cortex 12, 339–348

P1: S1364-6613(02)00008-6
http://tics.trends.com