nature human behaviour

LETTERS

DOI: 10.1038/s41562-017-0208-0

Meaning-based guidance of attention in scenes as revealed by meaning maps

John M. Henderson[1,2]* and Taylor R. Hayes[2]

Real-world scenes comprise a blooming, buzzing confusion of information. To manage this complexity, visual attention is guided to important scene regions in real time[1–7]. What factors guide attention within scenes? A leading theoretical position suggests that visual salience based on semantically uninterpreted image features plays the critical causal role in attentional guidance, with knowledge and meaning playing a secondary or modulatory role[8–11]. Here we propose instead that meaning plays the dominant role in guiding human attention through scenes. To test this proposal, we developed 'meaning maps' that represent the semantic richness of scene regions in a format that can be directly compared to image salience. We then contrasted the degree to which the spatial distributions of meaning and salience predict viewers' overt attention within scenes. The results showed that both meaning and salience predicted the distribution of attention, but that when the relationship between meaning and salience was controlled, only meaning accounted for unique variance in attention. This pattern of results was apparent from the very earliest time-point in scene viewing. We conclude that meaning is the driving force guiding attention through real-world scenes.

According to image guidance theories, attention is directed to scene regions on the basis of semantically uninterpreted image features. On this view, attention is, in a fundamental sense, a reaction to the image properties of the stimulus confronting the viewer, with attention 'pulled' to visually salient scene regions[12]. The most comprehensive theory of this type is based on visual salience, in which basic image features such as luminance contrast, colour and edge orientation are used to form a saliency map that provides the basis for attentional guidance[8,13,14].

An alternative theoretical perspective is represented by cognitive guidance theories, in which attention is directed to scene regions that are semantically informative. This position is consistent with strong evidence suggesting that humans are highly sensitive to the distribution of meaning in visual scenes from the earliest moments of viewing[7,15–17]. On this view, attention is primarily controlled by knowledge structures stored in memory that represent a scene. These knowledge structures contain information about a scene's likely semantic content and the spatial distribution of that content based on experience with general scene concepts and the specific scene instance currently in view[7]. On cognitive guidance theories, attention is 'pushed' to these meaningful scene regions by the cognitive system[2–7,18].

Most research on attentional guidance in scenes has focused on image salience. Little is currently known about how the spatial distribution of meaning across a scene influences attentional guidance. The emphasis on image salience is likely to be due in part to the relative ease of quantifying image properties and the relative difficulty of quantifying higher-level cognitive constructs related to scene meaning[4]. To test between image guidance and cognitive guidance theories, it is necessary to generate equivalent quantitative predictions from both meaning and salience that are in some sense on an equal footing.

Our central goal was to investigate the relative roles of meaning and salience in guiding attention through scenes. To capture the spatial distribution of meaning across a scene, we developed a method that represents scene meaning as a spatial map (a 'meaning map'). A meaning map can be taken as a conceptual analogue of a saliency map, capturing the distribution of semantic properties rather than image properties across a scene. Meaning maps can be directly compared to saliency maps and can also be used to predict attentional maps in the same manner as has been done with saliency maps[9,13,19,20]. With meaning maps in hand, we can directly compare the influences of meaning and salience on attentional guidance.

Meaning is spatially distributed in a non-uniform manner across a scene. Some scene regions are relatively rich in meaning, and others are relatively sparse. Here we generated meaning maps for scenes by asking subjects to rate the meaningfulness of scene regions. Digital photographs of real-world scenes (Fig. 1a) were divided into objectively defined and context-free circular overlapping regions at two spatial scales (Fig. 1b and c). Regions were presented independently of the scenes from which they were taken (Fig. 1d) and rated by naive raters on Mechanical Turk. We then built smoothed maps for each scene based on interpolated ratings over a large number of raters (Fig. 1e). (Details are given in Methods.)

It has been suggested that meaning and visual salience are likely to be highly correlated across scenes[3,18,21,22]. Yet this correlation has not so far been empirically tested. If such a correlation exists, then attentional effects that have been attributed to visual salience could be due to meaning[22–24]. Figure 2 presents the correlation of meaning and salience for each scene. On average, across the 40 scenes, the correlation was 0.80 (s.d. = 0.08). A one-sample t-test confirmed that the correlation was significantly greater than zero, $t(39) = 60.4$, $P < 0.0001$, 95% confidence interval (CI) [0.77, 0.82]. These findings establish that meaning and salience do indeed overlap substantially in scenes, as has previously been hypothesized. Meaning and salience also each accounted for unique variance (36% of the variance was not shared). To attribute attentional effects unambiguously to either meaning or salience, the effects of both must be considered together.

We can conceive of meaning maps and saliency maps as predictions concerning how attention will be guided through scenes. The empirical question is then how well the meaning and saliency

[1]Department of Psychology, University of California, Davis, CA 95618, USA. [2]Center for Mind and Brain, University of California, Davis, CA 95618, USA. *e-mail: johnhenderson@ucdavis.edu

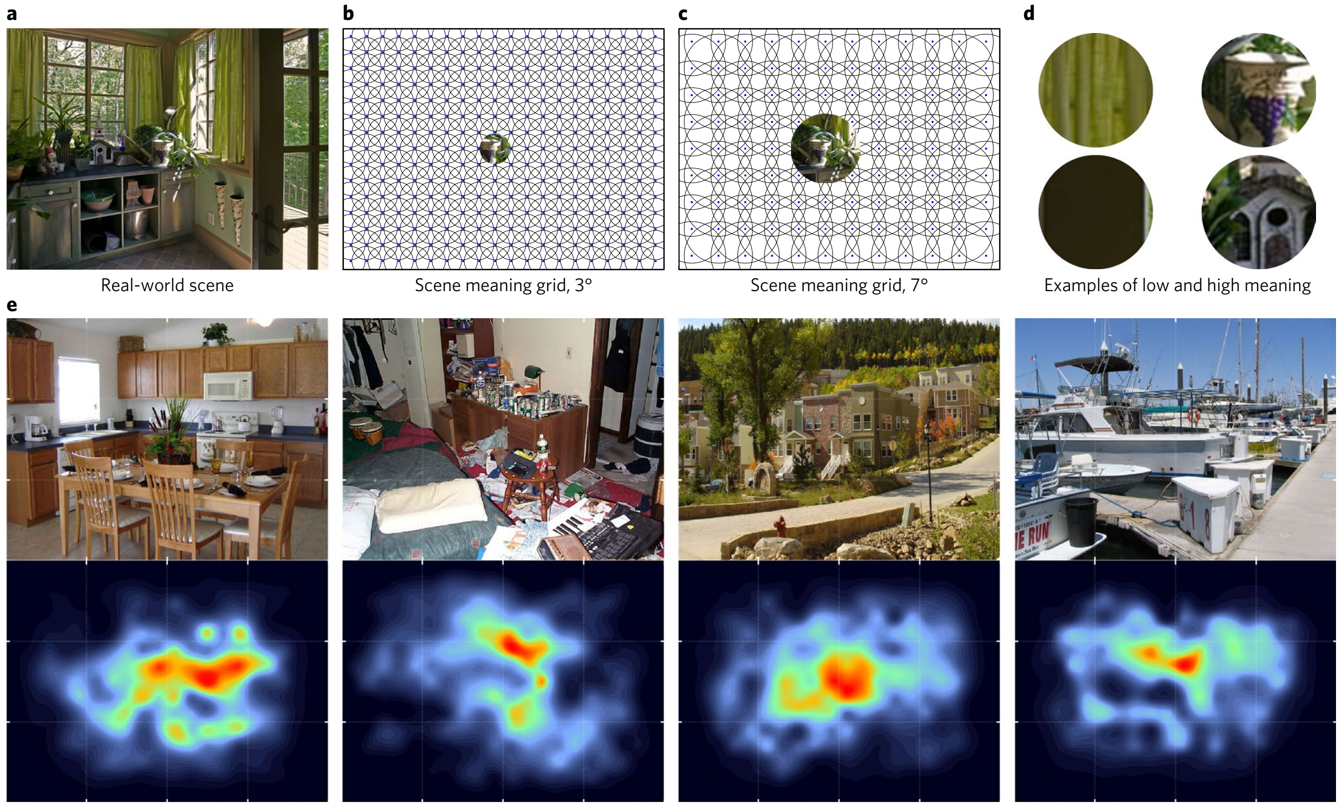

Fig. 1 | Generation of meaning maps. Meaning maps were generated from subject ratings ($N=165$) of context-free scene patches at two spatial scales. **a**–**c**, Each real-world scene (**a**) was decomposed into a series of overlapping circular patches at 3° (**b**) and 7° (**c**) spatial scales. Blue dots in **b** and **c** denote the centre of each circular patch that was rated, with example patches of the content captured by the 3° and 7° scales shown in the centre. **d**, Also shown are examples of high- and low-meaning patches. **e**, Ratings were combined to produce meaning maps, as shown for four example scenes.

maps predict observed distributions of attention. To answer this question, it is necessary to quantify attention over each scene. Following common practice in this literature, we operationalized the distribution of attention as the distribution of eye fixations. We had a group of human subjects view each scene for 12 seconds while their eye movements were recorded. Attention maps in the same format as the meaning and saliency maps were then generated from the eye movement data to represent where attention was directed (see Methods). Figure 3a shows a scene image with eye fixations superimposed, and Fig. 3b shows the attention map derived from these fixations.

Our next step was to determine how well the meaning maps (Fig. 3c) and saliency maps (Fig. 3d) predicted the spatial distribution of attention (Fig. 3a) as captured by attention maps (Fig. 3b). (Please see Supplementary Information for all scenes and their maps.) For this analysis, we used a method based on linear correlation to assess the degree to which meaning maps and saliency maps accounted for shared and unique variance in the attention maps[25].

Figure 4 presents the data for each of the 40 scenes using this approach. Each data point shows the R^2 value for the prediction maps (meaning and saliency) and the observed attention maps for saliency (blue) and meaning (red). Figure 4a shows the squared linear correlations. On average across the 40 scenes, meaning accounted for 53% of the variance in fixation density (mean $M=0.53$, s.d. $=0.11$) and saliency accounted for 38% of the variance in fixation density ($M=0.38$, s.d. $=0.12$). A two-tailed t-test revealed that this difference was statistically significant, $t(78)=5.63$, $P<0.0001$, 95% CI [0.10, 0.20].

To examine the unique variance in attention explained by meaning and salience when controlling for their shared variance, we computed squared semi-partial correlations. These correlations (Fig. 4b) revealed that across the 40 scenes, meaning captured more

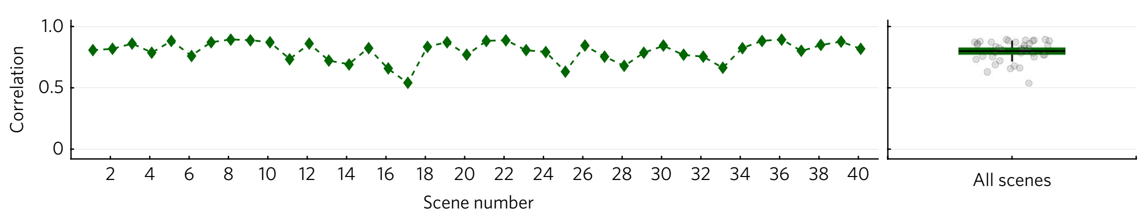

Fig. 2 | Correlation between saliency and meaning maps. The line plot shows the correlation between the meaning and saliency maps for each scene. The scatter box plot on the right shows the corresponding grand correlation mean across $N=40$ scenes (black horizontal line), 95% confidence intervals (coloured box) and 1 standard deviation (black vertical line). The mean correlation differed significantly from zero, $P<0.0001$.

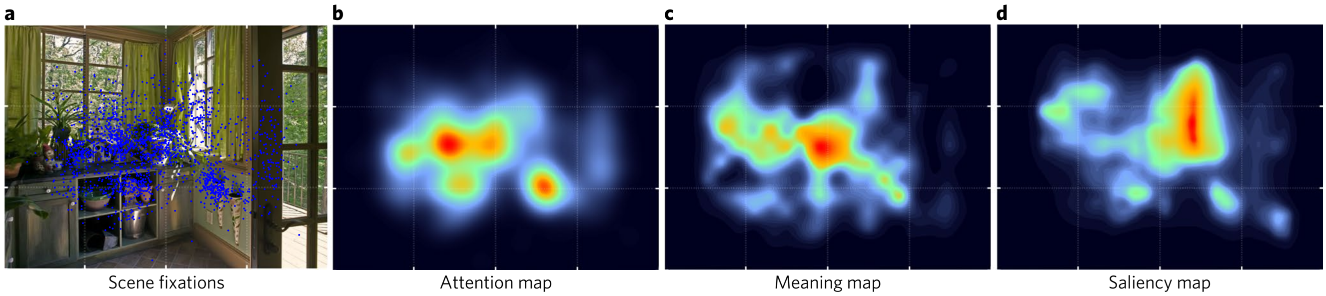

Fig. 3 | Attention, meaning and saliency maps for an example scene. a,**b**, We obtained eye movements (**a**) from subjects ($N=65$) who viewed each scene, and we generated attention maps (**b**) from those eye movement data. **c**,**d**, We compared the attention maps to the corresponding meaning maps (**c**) and saliency maps (**d**) from each scene.

than four times as much unique variance ($M=0.19$, s.d. $=0.10$) as saliency ($M=0.04$, s.d. $=0.04$). Meaning maps accounted for a statistically significant 19% additional variance in the attention maps after controlling for salience, whereas saliency maps accounted for a non-significant 4% additional variance after controlling for meaning. A two-tailed t-test confirmed that this difference was statistically significant, $t(78)=8.42$, $P<0.0001$, 95% CI [0.11, 0.18]. Additional analyses indicated that these results held when the scene centres were removed from the analysis, suggesting that they were not due to a concentration of attention at the centres of the scenes, and they were also replicated when using a task involving free-viewing for aesthetic judgement, suggesting that they were not an artifact of the memorization viewing task (see Supplementary Information). Overall, the results showed that meaning was better able than salience to explain the distribution of attention over scenes.

So far, we have examined the roles of meaning and salience over the entire viewing period for each scene. However, it has been proposed that attention is initially guided by image salience, but that over time, as knowledge representations become available and meaning can be acquired from more of the scene, meaning begins to play a greater role[7,26,27].

To investigate whether the effects of meaning and salience changed over time as each scene was viewed, we conducted

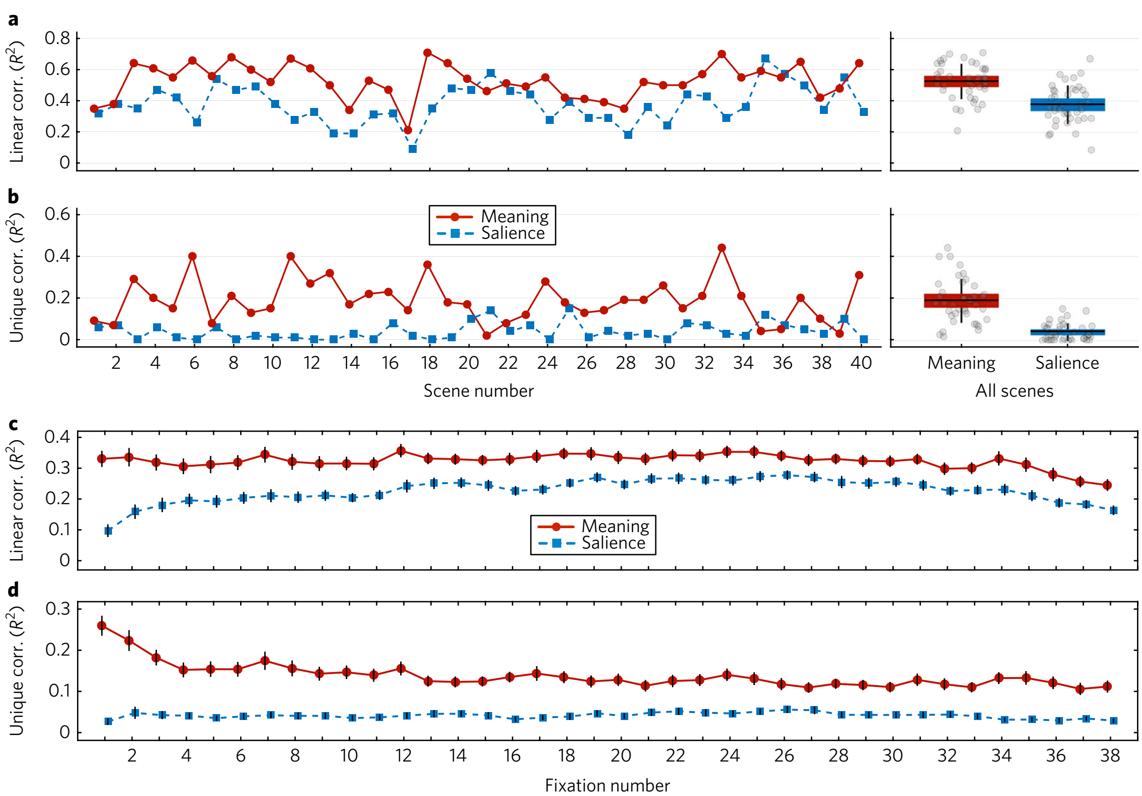

Fig. 4 | Squared linear correlation and semi-partial correlation by scene and by fixation order. a,**b**, Shown for each scene are the linear correlation (**a**) and semi-partial correlation (**b**), between fixation density and meaning (red) and fixation density and salience (blue). The scatter box plots on the right show the corresponding grand correlation means across $N=40$ scenes (black horizontal line), 95% confidence intervals (coloured box) and 1 standard deviation (black vertical line). Both linear and semi-partial correlations for meaning and salience differed significantly, $P<0.0001$. **c**,**d**, Plots also show the squared linear correlation (**c**) and corresponding semi-partial correlation (**d**), between fixation density and meaning (red) and fixation density and salience (blue), as a function of fixation order across all 40 scenes. Error bars represent standard error of the mean. Correlations (corr.) differed significantly at all fixations, FDR $P<0.05$.

temporal time-step analyses. Linear correlation and semi-partial correlation were computed as described above, but were based on a series of attention maps generated from each sequential eye fixation (first, second, third and so on) in each scene. The results are shown in Fig. 4. For the linear correlations, the relationship was stronger between meaning and fixation maps for all time steps (Fig. 4c) and was very consistent across the 40 scenes. Meaning accounted for 33.0%, 33.6% and 31.9% of the variance in the first three fixations, whereas salience accounted for only 9.7%, 15.9% and 18.1% of the variance in the first three fixations, respectively. Two-sample, two-tailed t-tests were performed for all 38 time points, and p-values were corrected for multiple comparisons using the false discovery rate (FDR) correction[28]. This procedure confirmed the advantage for meaning over salience at all 38 time points (FDR $P < 0.05$).

When controlling for the correlation among the two prediction maps with semi-partial correlations, the advantage for the meaning maps observed in the overall analyses was also found to hold across time steps (Fig. 4d). The same testing and false discovery rate correction revealed that all 38 time points were significantly different (FDR $P < 0.05$), with meaning accounting for 25.9%, 22.4% and 18.2% of the unique variance in the first three fixations, whereas salience accounted for 2.7%, 4.8% and 4.2% of the unique variance in the first three fixations, respectively. In sum, counter to the salience-first hypothesis, we observed no crossover of the effects of meaning and salience over time. Instead, in both the correlation and semi-partial correlation analyses, we observed an advantage for meaning from the very first fixation. Indeed, if anything, there was an even greater advantage for meaning in guiding attention over the first few fixations than later in viewing. These results indicate that meaning begins to guide attention as soon as a scene appears, consistent with past findings that viewing task can also override salience as soon as the first saccade[23,29].

The dominant role of meaning in guiding attention can be accommodated by a theoretical perspective that places explanatory primacy on scene semantics. For example, according to the cognitive relevance model[22,23], the role of a particular object or scene region in guiding attention is determined solely by its meaning in the context of the scene and the current goals of the viewer, and not by its visual salience. In this view, meaning determines attentional priority, with image properties used to provide perceptual objects and regions to which attentional priority can be assigned based on knowledge representations. In this model, the visual stimulus is used to generate the perceptual objects and other potential targets for attention, but the image features themselves provide a flat (that is, unranked) landscape of attentional targets, with attentional priority rankings provided by knowledge representations[3,22,23]. Note that in this view the meaning of all objects and scene regions across the entire scene need not be established during the initial glimpse. Rather, rapidly ascertained scene gist[7,30–32] can be used to generate predictions about what objects are likely to be informative and where those objects are likely to be found[4]. This knowledge combined with representations of perceptual objects generated from peripheral visual information would be sufficient to guide attention using meaning. In addition, given that saccade amplitudes tend to be relatively short in scene viewing (about 3.5° on average in the present study), meaning directly acquired from parafoveal scene regions during each fixation would often be available to guide the next attentional shift to a meaningful region.

In summary, we found that meaning was better able than visual salience to account for the guidance of attention through real-world scenes. Furthermore, we found that the influence of meaning was apparent both at the very beginning of scene viewing and throughout the viewing period. Given the strong correlation between meaning and salience observed here, and the fact that only meaning accounted for unique variance in the distribution of attention, we conclude that both previous and current results are consistent with a theory in which meaning is the dominant force guiding attention through scenes. This conclusion has important implications for current theories of attention across diverse disciplines that have been influenced by image salience theory, including vision science, cognitive science, visual neuroscience and computer vision.

Methods

Meaning maps. *Subjects.* Scene patches were rated by 165 subjects on Amazon Mechanical Turk. Subjects were recruited from the United States, had a HIT (human intelligence task) approval rate of 99% and 500 HITs approved, and were only allowed to participate in the study once. Subjects were paid \$0.50 cents per assignment, and all subjects provided informed consent.

Stimuli. Stimuli consisted of 40 digitized photographs of real-world scenes. Each scene was decomposed into a series of partially overlapping and tiled circular patches at two spatial scales of 3° and 7° (Fig. 1). Simulated recovery of known scene properties (such as luminance) indicated that the underlying known property could be recovered well (98% variance explained) using these two spatial scales with patch overlap. The full patch stimulus set consisted of 12,000 unique 3° patches and 4,320 unique 7° patches for a total of 16,320 scene patches.

Procedure. Each subject rated 300 random scene patches extracted from 40 scenes. Subjects were instructed to assess the meaningfulness of each patch based on how informative or recognizable they thought it was. Subjects were first given examples of two low-meaning and two high-meaning scene patches to make sure they understood the rating task. Subjects then rated the meaningfulness of test patches on a six-point Likert scale ('very low', 'low', 'somewhat low', 'somewhat high', 'high', 'very high'). Patches were presented in random order and without scene context, so ratings were based on context-independent judgments. Each unique patch was rated three times by three independent raters for a total of 48,960 ratings. However, owing to the high degree of overlap across patches, each 3° patch contained rating information from 27 independent raters, and each 7° patch from 63 independent raters.

Meaning maps were generated from the ratings by averaging, smoothing and combining 3° and 7° maps from the corresponding patch ratings. The ratings for each pixel at each scale (3° and 7°) in each scene were averaged, producing an average 3° and 7° rating map for each scene. The average 3° and 7° rating maps were then smoothed using thin-plate spline interpolation (MATLAB 'fit' using the 'thinplateinterp' method). Finally, the smoothed 3° and 7° maps were combined using a simple average: (3° map + 7° map)/2. This procedure was used to create a meaning map for each scene. The final map was blurred using a Gaussian kernel followed by a multiplicative centre bias operation which down-weighted the periphery to account for the central fixation bias, the commonly observed phenomenon in which subjects concentrate their fixations more centrally and rarely fixate the outside border of a scene[33]. This centre bias operation is also commonly applied to saliency maps.

Saliency maps. To investigate the relationship between the generated meaning maps and image-based saliency maps, saliency maps for each scene were computed using the Graph-based Visual Saliency (GBVS) toolbox with default settings[14]. GBVS is a prominent saliency model that combines conspicuity maps of different low-level image features. The same centre bias operation described for the meaning maps was applied to the saliency maps to down-weight the periphery.

Histogram matching. The meaning and saliency maps were normalized to a common scale using image histrogram matching, with the attention map for each scene serving as the reference image for the corresponding meaning and saliency maps. Histogram matching of the meaning and saliency maps was accomplished using the MATLAB function 'imhistmatch' in the Image Processing Toolbox.

Eyetracking experiment and attention maps. *Subjects.* Seventy-nine University of South Carolina undergraduate students with normal or corrected-to-normal vision participated in the experiment. All subjects were naive concerning the purposes of the experiment and provided informed consent as approved by the University of South Carolina Institutional Review Board. In MATLAB, the eye movement data from each subject were inspected for excessive artifacts caused by blinks or loss of calibration due to incidental movement by examining the mean percentage of signal across all trials. Fourteen subjects with less than 75% signal were removed, leaving 65 subjects that were tracked very well (mean signal 91.74%).

Apparatus. Eye movements were recorded with an EyeLink 1000+ tower mount eyetracker (spatial resolution 0.01) sampling at 1,000 Hz. Subjects sat 90 cm away from a 21′′ monitor, so that scenes subtended approximately 33° × 25° of visual angle. Head movements were minimized by using a chin and forehead rest. Although viewing was binocular, eye movements were recorded from the right eye. The experiment was controlled with SR Research Experiment Builder software.

Stimuli. Stimuli consisted of the same 40 digitized photographs of real-world scenes that were used to create the meaning and saliency maps.

Procedure. Subjects were instructed to memorize each scene in preparation for a later memory test. The memory test was not administered. Each trial began with fixation on a cross at the centre of the display for 300 ms. Following central fixation, each scene was presented for 12 seconds while eye movements were recorded. Scenes were presented in the same order across all 79 subjects.

A 13-point calibration procedure was performed at the start of each session to map eye position to screen coordinates. Successful calibration required an average error of less than 0.49° and a maximum error of less than 0.99°. Fixations and saccades were segmented with EyeLink's standard algorithm using velocity and acceleration thresholds ($30° s^{-1}$ and $9,500° s^{-1}$).

Eye movement data were imported offline into MATLAB using the EDFConverter tool. In MATLAB, the eye movement data from each participant were inspected for excessive artifacts caused by blinks or loss of calibration due to incidental movement by examining the mean percentage of signal across all trials. The first fixation, always located at the centre of the display as a result of the pretrial fixation period, was discarded.

Attention maps. Across subjects, for every position (that is, each *x,y* coordinate pair) within a scene, +1 was accumulated for each fixation, producing a fixation frequency matrix. A Gaussian low-pass filter with circular boundary conditions and a cutoff frequency of −6 dB was applied to the fixation frequency matrix for each scene to account for foveal acuity and eye tracker error.

Data availability. Scene images, meaning maps, saliency maps and attention maps for all scene stimuli are shown in the Supplementary Information. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 3 May 2017; Accepted: 18 August 2017;
Published online: 25 September 2017

References

1. Land, M. F. & Hayhoe, M. M. In what ways do eye movements contribute to everyday activities? *Vision Res.* **41**, 3559–3565 (2001).
2. Hayhoe, M. M. & Ballard, D. Eye movements in natural behavior. *Trends Cogn. Sci.* **9**, 188–194 (2005).
3. Henderson, J. M. Human gaze control during real-world scene perception. *Trends Cogn. Sci.* **7**, 498–504 (2003).
4. Henderson, J. M. Gaze control as prediction. *Trends Cogn. Sci.* **21**, 15–23 (2017).
5. Buswell, G. T. *How People Look at Pictures* (Univ. Chicago Press, Chicago, 1935).
6. Yarbus, A. L. *Eye Movements and Vision* (Plenum, 1967).
7. Henderson, J. M. & Hollingworth, A. High-level scene perception. *Annu. Rev. Psychol.* **50**, 243–271 (1999).
8. Itti, L. & Koch, C. Computational modelling of visual attention. *Nat. Rev. Neurosci.* **2**, 194–203 (2001).
9. Parkhurst, D., Law, K. & Niebur, E. Modelling the role of salience in the allocation of visual selective attention. *Vision Res.* **42**, 107–123 (2002).
10. Borji, A., Sihite, D. N. & Itti, L. Quantitative analysis of human-model agreement in visual saliency modeling: a comparative study. *IEEE Trans. Image Process.* **22**, 55–69 (2013).
11. Walther, D. & Koch, C. Modeling attention to salient proto-objects. *Neural Networks* **19**, 1395–1407 (2006).
12. Henderson, J. M. Regarding scenes. *Curr. Dir. Psychol. Sci.* **16**, 219–222 (2007).
13. Itti, L., Koch, C. & Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell* **20**, 1254–1259 (1998).
14. Harel, J., Koch, C. & Perona, P. in *Advances in Neural Information Processing Systems (NIPS 2006)* Vol. 19, 1–8 (2006).
15. Potter, M. Meaning in visual search. *Science* **187**, 965–966 (1975).
16. Biederman, I. Perceiving real-world scenes. *Science* **177**, 77–80 (1972).
17. Wolfe, J. M. & Horowitz, T. S. Five factors that guide attention in visual search. *Nat. Hum. Behav.* **1**, 0058 (2017).
18. Tatler, B. W., Hayhoe, M. M., Land, M. F. & Ballard, D. H. Eye guidance in natural vision: reinterpreting salience. *J. Vis.* **11**, 5 (2011).
19. Carmi, R. & Itti, L. The role of memory in guiding attention during natural vision. *J. Vis.* **6**, 898–914 (2006).
20. Torralba, A., Oliva, A., Castelhano, M. S. & Henderson, J. M. Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychol. Rev.* **113**, 766–786 (2006).
21. Elazary, L. & Itti, L. Interesting objects are visually salient. *J. Vis.* **8**, 3.1–15 (2008).
22. Henderson, J. M., Brockmole, J. R., Castelhano, M. S. & Mack, M. in *Eye Movements: A Window on Mind and Brain* (eds. Van Gompel, R. P. G. et al.) 537–562 (Elsevier, Oxford, 2007).
23. Henderson, J. M., Malcolm, G. L. & Schandl, C. Searching in the dark: cognitive relevance drives attention in real-world scenes. *Psychon. Bull. Rev.* **16**, 850–856 (2009).
24. Nuthmann, A. & Henderson, J. M. Object-based attentional selection in scene viewing. *J. Vis.* **10**, 20 (2010).
25. Bylinskii, Z., Judd, T., Oliva, A., Torralba, A. & Durand, F. What do different evaluation metrics tell us about saliency models? Preprint at http://arxiv.org/abs/1604.03605 (2016).
26. Anderson, N. C., Donk, M. & Meeter, M. The influence of a scene preview on eye movement behavior in natural scenes. *Psychon. Bull. Rev*. **23**, 1794–1801 (2016).
27. Anderson, N. C., Ort, E., Kruijne, W., Meeter, M. & Donk, M. It depends on when you look at it: salience influences eye movements in natural scene viewing and search early in time. *J. Vis.* **15**, 9 (2015).
28. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).
29. Einhäuser, W., Rutishauser, U. & Koch, C. Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *J. Vis.* **8**, 2.1–19 (2008).
30. Oliva, A. & Torralba, A. Building the gist of a scene: the role of global image features in recognition. *Prog. Brain Res.* **155**, 23–36 (2006).
31. Castelhano, M. S. & Henderson, J. M. The influence of color on the perception of scene gist. *J. Exp. Psychol. Hum. Percept. Perform* **34**, 660–675 (2008).
32. Castelhano, M. S. & Henderson, J. M. Flashing scenes and moving windows: an effect of initial scene gist on eye movements. *J. Vis.* **3**, 67a (2003).
33. Tatler, B. W. The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor biases and image feature distributions. *J. Vis.* **7**, 4.1–17 (2007).

Acknowledgements

We thank the members of the UC Davis Visual Cognition Research Group for their feedback and comments. This research was partially funded by BCS-1636586 from the US National Science Foundation. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

J.M.H. conceived of and designed the study, and drafted and revised the manuscript. T.R.H. designed the study, collected and analysed the data, and revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at doi:10.1038/s41562-017-0208-0.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to J.M.H.

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

nature research

Corresponding author(s): John Henderson

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Sample size was based on our prior research in this area. Sample size and stopping rule were set a priori.

2. Data exclusions

Describe any data exclusions.

The exclusion criteria are given in the Methods and are based on published standards in the eye movement literature. They were set a priori.

3. Replication

Describe whether the experimental findings were reliably reproduced.

The experimental findings were reproduced, and the replication is reported in the Supplemental Materials.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

N/A. All subjects saw all stimuli in the same order.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

N/A

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

June 2017

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	SR Research Data Viewer and Matlab functions were used to analyze the data. The specific Matlab functions that were used are named in the methods.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	N/A

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	N/A

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	N/A
b. Describe the method of cell line authentication used.	N/A
c. Report whether the cell lines were tested for mycoplasma contamination.	N/A
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	N/A

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	N/A

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	Participants were University of South Carolina undergraduates, all subjects were naive concerning the purposes of the experiment, and all subjects provided signed informed consent as approved by the University of South Carolina Institutional Review Board.