

In the format provided by the authors and unedited.

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

John M. Henderson^{1,2*} and Taylor R. Hayes²

¹Department of Psychology, University of California, Davis, CA 95618, USA. ²Center for Mind and Brain, University of California, Davis, CA 95618, USA.

*e-mail: johnhenderson@ucdavis.edu

Meaning-Based Guidance of Attention in Scenes as Revealed by Meaning Maps

John M. Henderson^{1,2} and Taylor R. Hayes²

¹Department of Psychology, University of California, Davis

²Center for Mind and Brain, University of California, Davis

Supplementary Methods

Analyses Excluding Map Centers

To ensure that the advantage of meaning maps over saliency maps in predicting attention was not due to a center bias advantage for the meaning maps, we conducted a supplementary set of analyses in which the data from the central 3 degrees of each map was removed from consideration. Differences in map successes in this analysis could therefore not be attributed to differences in the ability of meaning maps to predict central fixations. The results of these analyses was qualitatively and quantitatively very similar to the complete analyses.

Supplementary Figure 1 shows the correlation of meaning and salience for each scene. On average across all scenes the correlation was 0.81 (SD=0.07). A one sample t-test confirmed that the correlation was significantly greater than zero, $t(39) = 69.5, p < .0001$, 95% CI [0.79, 0.84]. Meaning and salience also each accounted for unique variance (i.e., 34% of the variance was not shared).

Supplementary Figure 2 presents the linear correlation data used to assess the degree to which meaning maps and saliency maps accounted for shared and unique variance in the attention maps data for each scene. 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). The top of Supplementary Figure 2 shows the squared linear correlations. On average across the 40 scenes, meaning accounted for 52% of the variance in fixation density (M=0.52, SD=0.11) and saliency account for 37% of the variance in fixation density (M = 0.37, SD = 0.13). A two-tailed t-test revealed this difference was statistically significant, $t(78) = 5.53, p < .0001$, 95% CI [0.09, 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, shown in the bottom of Supplementary Figure 2, revealed that across the 40 scenes, meaning captured more than 4 times as much unique variance (M=0.19, SD=0.10)

as saliency ($M = 0.04$, $SD = 0.04$). A two-tailed t-test confirmed that this difference was statistically significant, $t(78) = 8.27$, $p < .0001$, 95% CI [0.11, 0.18]. These results confirm those of the complete analysis and indicate that meaning was better able than salience to explain the distribution of attention over scenes even when scene centers were excluded.

Supplementary Figure 3 shows the temporal time-step analyses with the map centers removed. Linear correlation and semi-partial correlation were conducted as in the main time-step analyses based on a series of attention maps generated from each sequential eye fixation (1st, 2nd, 3rd, etc.) in each scene. Using the same testing and false discovery rate correction as in the main analyses, all 38 time points were significantly different in both the linear and semi-partial analyses ($FDR < 0.05$). In the linear correlation analysis (top of Supplementary Figure 3), meaning accounted for 29.8%, 32.2%, and 31.0% of the variance in the first 3 fixations, whereas salience accounted for only 9.6%, 15.7%, and 17.6% of the variance in the first 3 fixations. 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, as shown in the bottom of Supplementary Figure 3. Meaning accounting for 23.3%, 21.6%, and 17.8% of the unique variance in the first 3 fixations, whereas salience accounted for 3.3%, 5.0%, and 4.4% of the unique variance in the first 3 fixations, respectively.

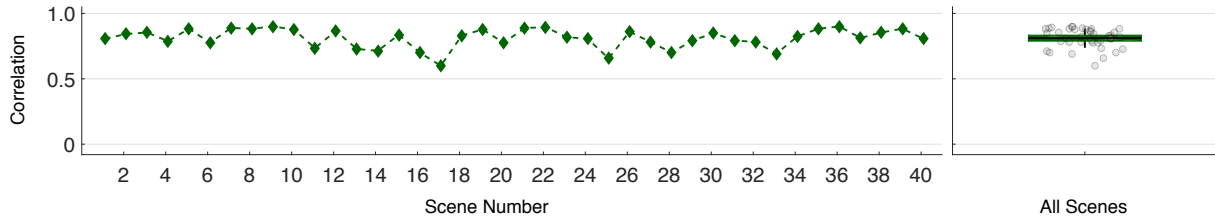
Replication with Aesthetic Judgment Task

To ensure that the observed results replicate over viewing instruction, we ran a second experiment using twelve of the original scenes under two viewing instructions. The twelve scenes were selected at random before the results of the main experiment were known. Subjects were instructed to memorize the scenes (memorization task) or to indicate how much they liked each scene on a 1-3 scale (aesthetic judgment task). Forty-six subjects viewed the twelve scenes, with each subject seeing six scenes in each of the instruction conditions. Each subject saw each scene once, with assignment of scene to task counterbalanced across subjects, so data for each scene in each condition was based on 23 subjects. Order of task

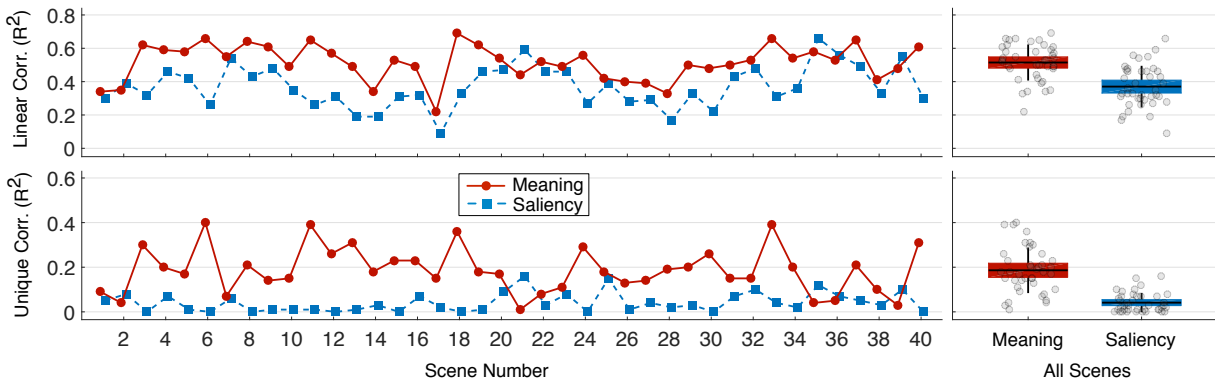
was counterbalanced across subjects. Scenes were each viewed for twelve seconds as in the main experiment. Attention maps generated from these subjects were then compared to the meaning and saliency maps as described in the main experiment. The results for the memorization and aesthetic judgment tasks are shown in Supplementary Figures 4 and 5 respectively. The new data in each figure are shown as individual data points superimposed on the data figures from the original experiment. Each data point shows the R^2 value for each prediction map (meaning and saliency) and the observed attention maps. The top panels of Supplementary Figures 4 and 5 show the squared linear correlations. For the memorization task, on average across the 12 scenes, meaning accounted for 52% of the variance in fixation density ($M=0.52$, $SD=0.12$) and saliency account for 30% of the variance in fixation density ($M=0.30$, $SD=0.12$). A two-tailed ttest revealed this difference was statistically significant, $t(22) = 4.57, p < .0001$, 95% CI [0.12, 0.33]. For the aesthetic judgment task, on average across the 12 scenes, meaning accounted for 57% of the variance in fixation density ($M=0.57$, $SD=0.09$) and saliency account for 30% of the variance in fixation density ($M=0.30$, $SD=0.12$). A two-tailed t-test revealed this difference was statistically significant, $t(22) = 6.27, p < .0001$, 95% CI [0.18, 0.37].

The bottom panels of Supplementary Figures 4 and 5 show the squared semi-partial correlations to examine the unique variance in attention explained by meaning and salience when controlling for their shared variance. For the memorization task, these correlations revealed that across the 12 scenes, meaning captured more than 20 times as much unique variance ($M=0.24$, $SD=0.10$) as saliency ($M=0.01$, $SD=0.02$). A two-tailed t-test revealed this difference was statistically significant, $t(22) = 7.50, p < .0001$, 95% CI [0.16, 0.28]. For the aesthetic judgment task, across the 12 scenes meaning captured more than 25 times as much unique variance ($M=0.29$, $SD=0.11$) as saliency ($M=0.01$, $SD=0.02$). A two-tailed t-test revealed this difference was statistically significant, $t(22) = 8.35, p < .0001$, 95% CI [0.21, 0.34]. These results confirm those of the main experiment and indicate that meaning was better able than salience to explain the distribution of attention over scenes even when scene centers were excluded.

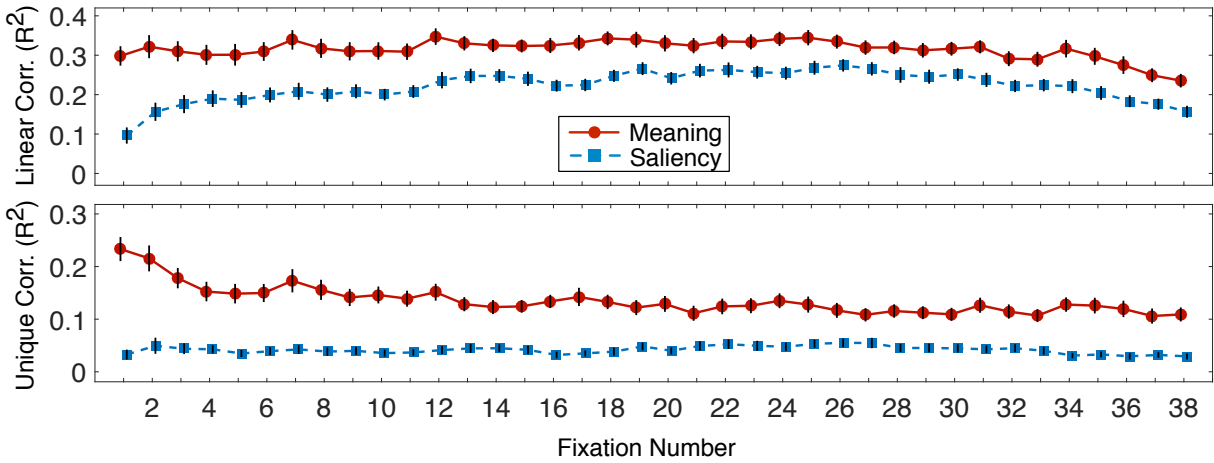
Supplementary Figures



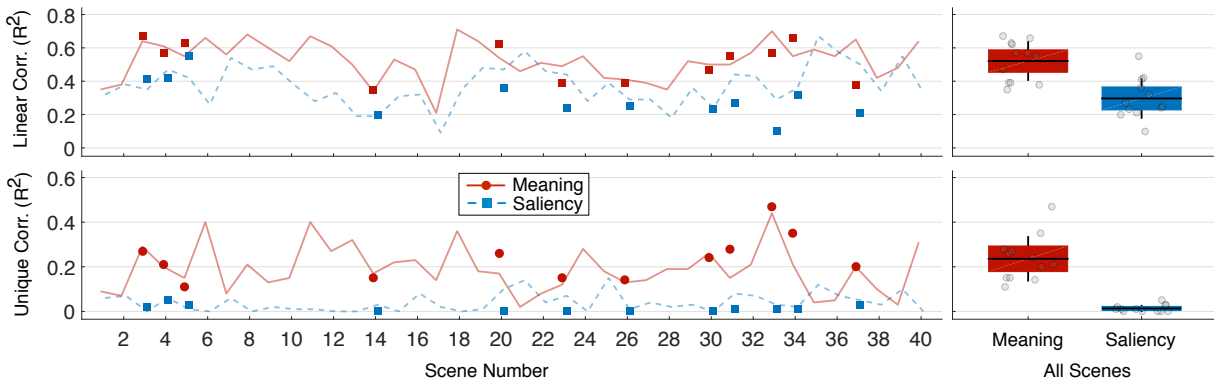
Supplementary Figure 1 *Correlation between saliency and meaning maps excluding map centers.* The line plot shows the correlation between saliency and meaning maps for each scene. The scatter box plot on the right shows the corresponding grand mean (black horizontal line), 95% confidence intervals (colored box), and 1 standard deviation (black vertical line) across all 40 scenes.



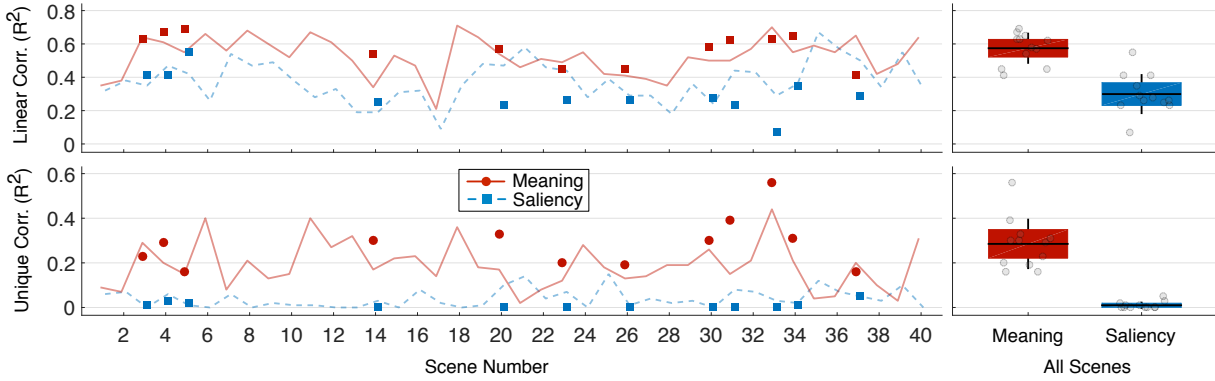
Supplementary Figure 2 *Squared linear correlation and semi-partial correlation by scene and across all scenes excluding map centers.* The line plots show the linear correlation (top) and semi-partial correlation (bottom) between fixation density and meaning (red) and saliency (blue) by scene. The scatter box plots on the right show the corresponding grand mean (black horizontal line), 95% confidence intervals (colored box), and 1 standard deviation (black vertical line) for meaning and saliency across all 40 scenes.



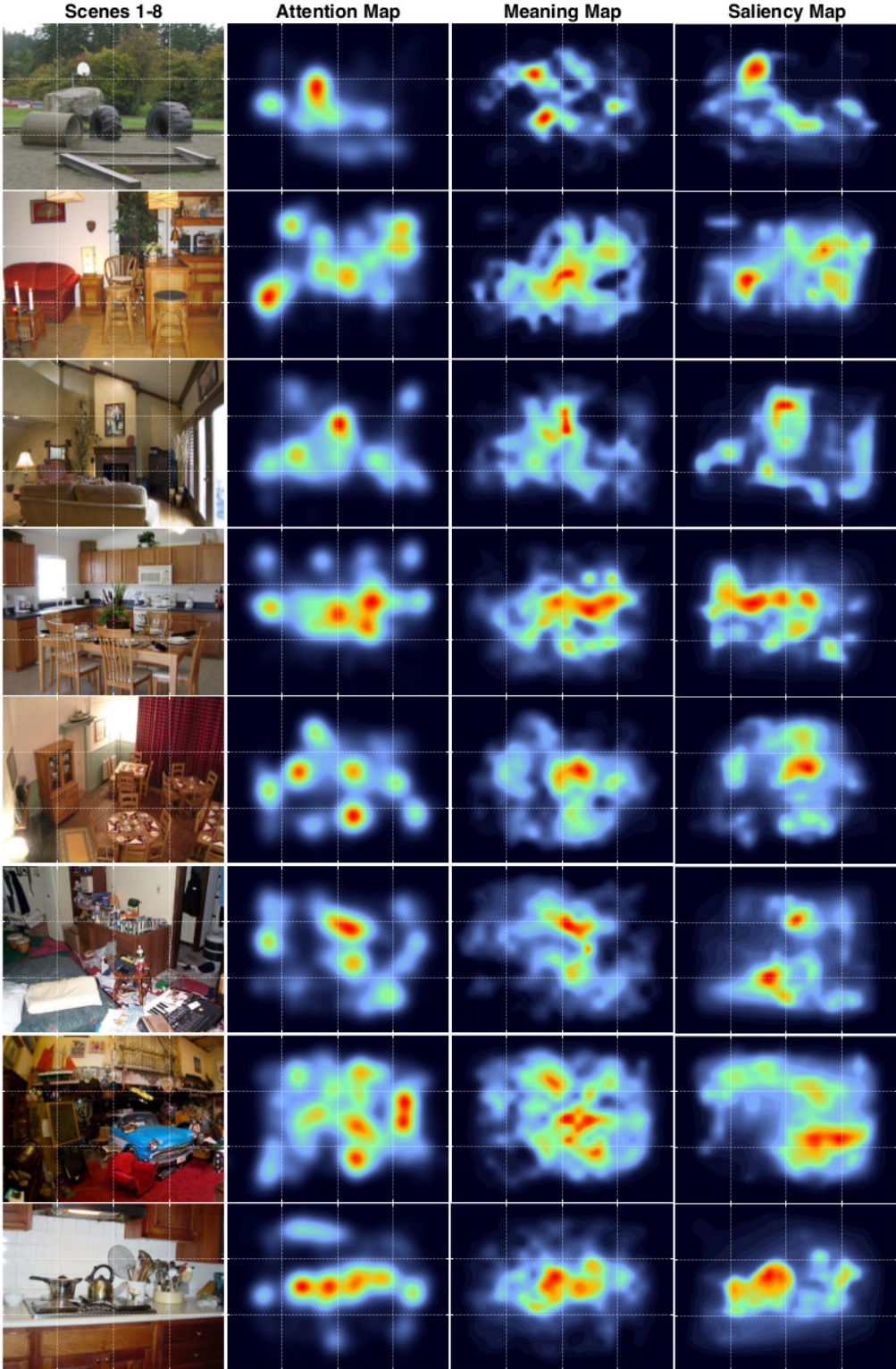
Supplementary Figure 3 Squared linear correlation and squared semi-partial correlation as a function of fixation number excluding map centers. The top panel shows the squared linear correlation between fixation density and meaning (red) and saliency (blue) as a function of fixation order across all 40 scenes. The bottom panel shows the corresponding semi-partial correlation as a function of fixation order across all 40 scenes. Error bars represent standard error of the mean.



Supplementary Figure 4 Squared linear correlation and semi-partial correlation by scene and across scenes for the memorization condition of the replication experiment (dark symbols) superimposed on the data from the original experiment (light lines). The plots show the linear correlation (top) and semi-partial correlation (bottom) between fixation density and meaning (red) and saliency (blue) by scene. The scatter box plots on the right show the corresponding grand mean (black horizontal line), 95% confidence intervals (colored box), and 1 standard deviation (black vertical line) for meaning and saliency across all 12 scenes.



Supplementary Figure 5 Squared linear correlation and semi-partial correlation by scene and across scenes for the aesthetic judgment condition of the replication experiment (dark symbols) superimposed on the data from the original experiment (light lines). The plots show the linear correlation (top) and semi-partial correlation (bottom) between fixation density and meaning (red) and saliency (blue) by scene. The scatter box plots on the right show the corresponding grand mean (black horizontal line), 95% confidence intervals (colored box), and 1 standard deviation (black vertical line) for meaning and saliency across all 12 scenes.



Scenes 9-16

Attention Map

Meaning Map

Saliency Map

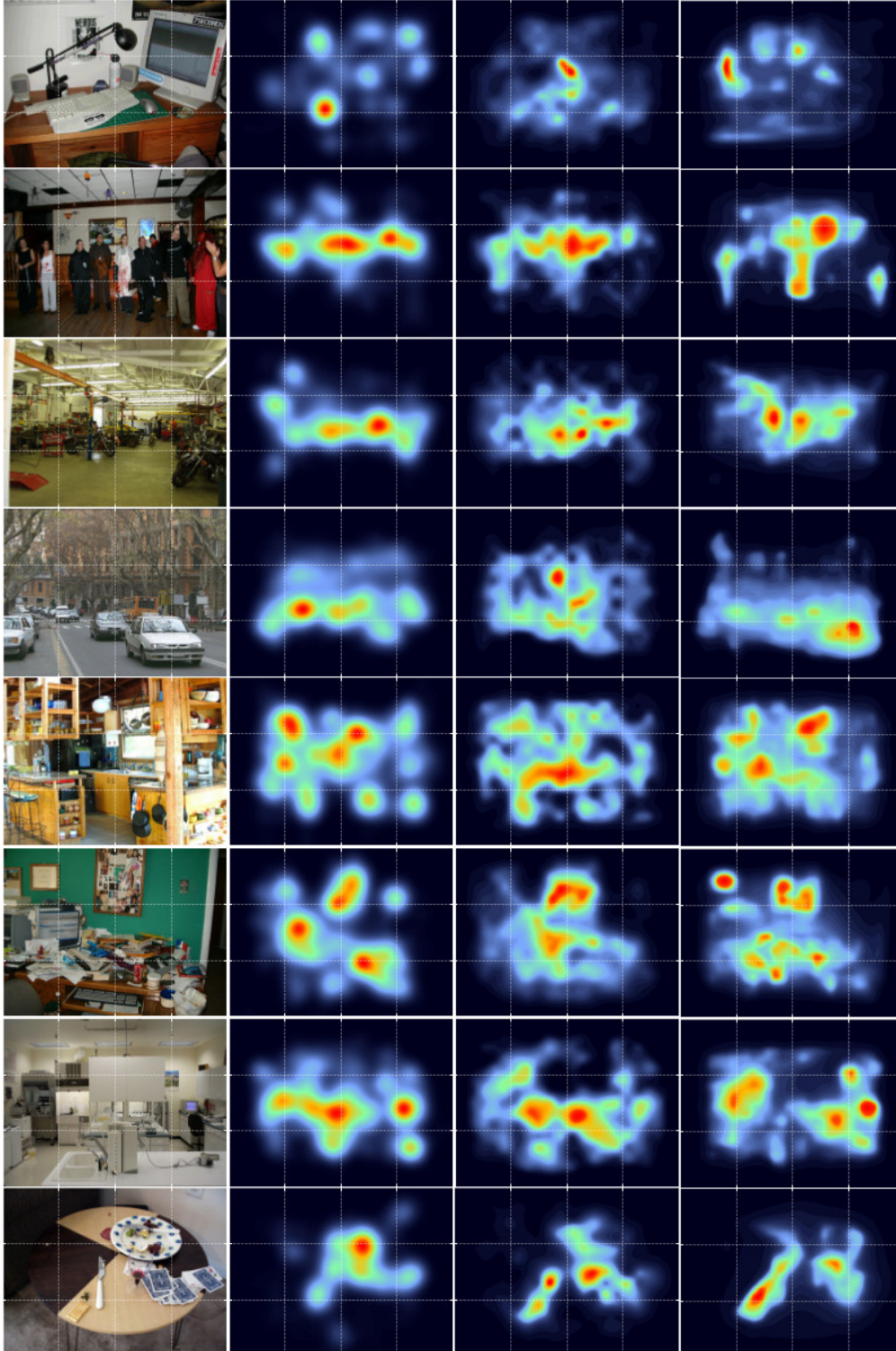


Scenes 17-24

Attention Map

Meaning Map

Saliency Map

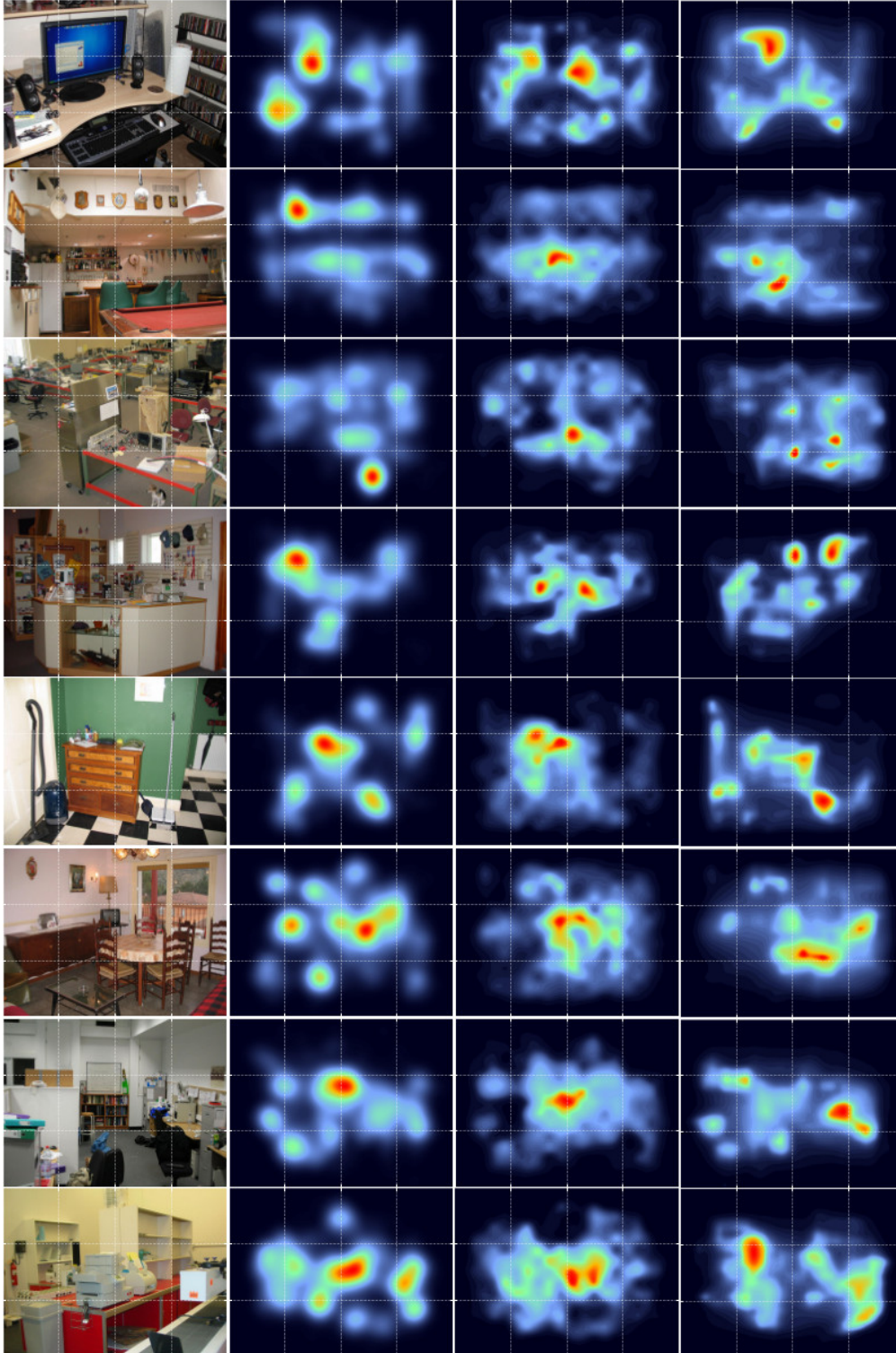


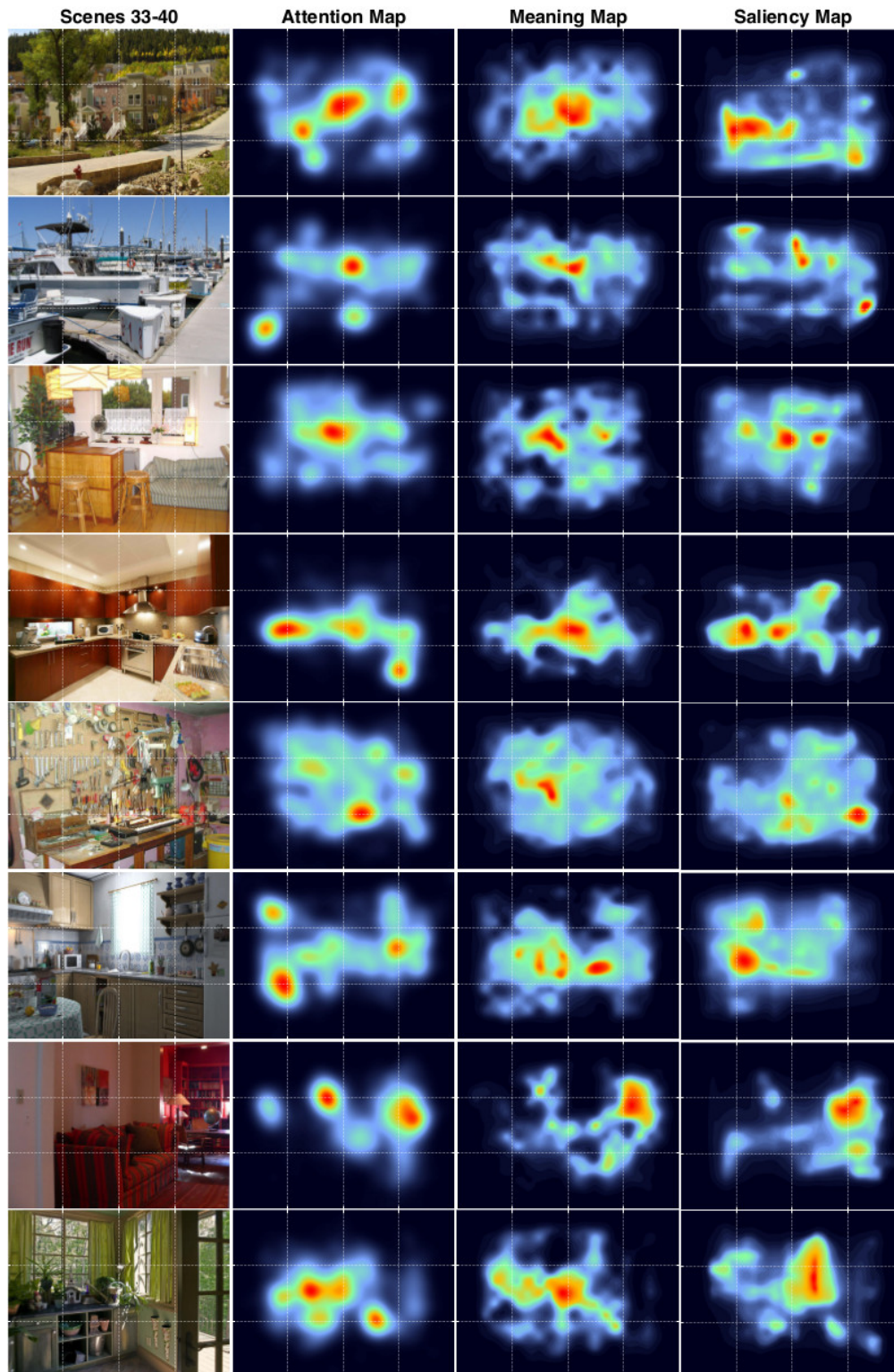
Scenes 25-32

Attention Map

Meaning Map

Saliency Map





Supplementary Figure 6 All scenes and corresponding attention, meaning, and saliency maps. Figure by Henderson & Hayes, 2017; available at <http://dx.doi.org/10.6084/m9.figshare.5306572> under a CC-BY4.0 license.