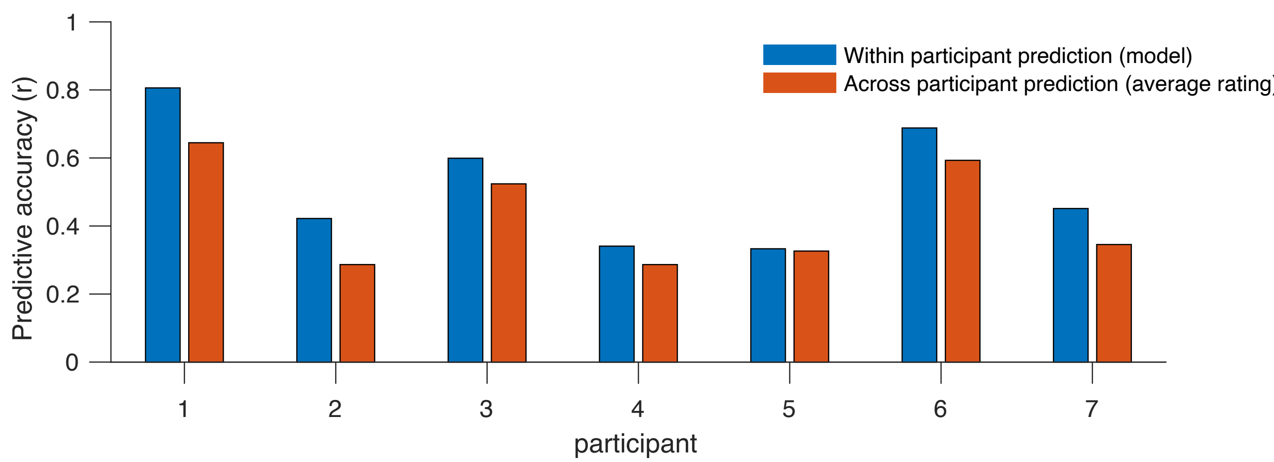
Supplementary materials for “Aesthetic preference for art can be predicted from a mixture of low- and high-level visual features”

Kiyohito Iigaya *[1], Sanghyun Yi[1], Iman A. Wahle[1], Koranis Tanwisuth[1], and John P. O’Doherty[1]

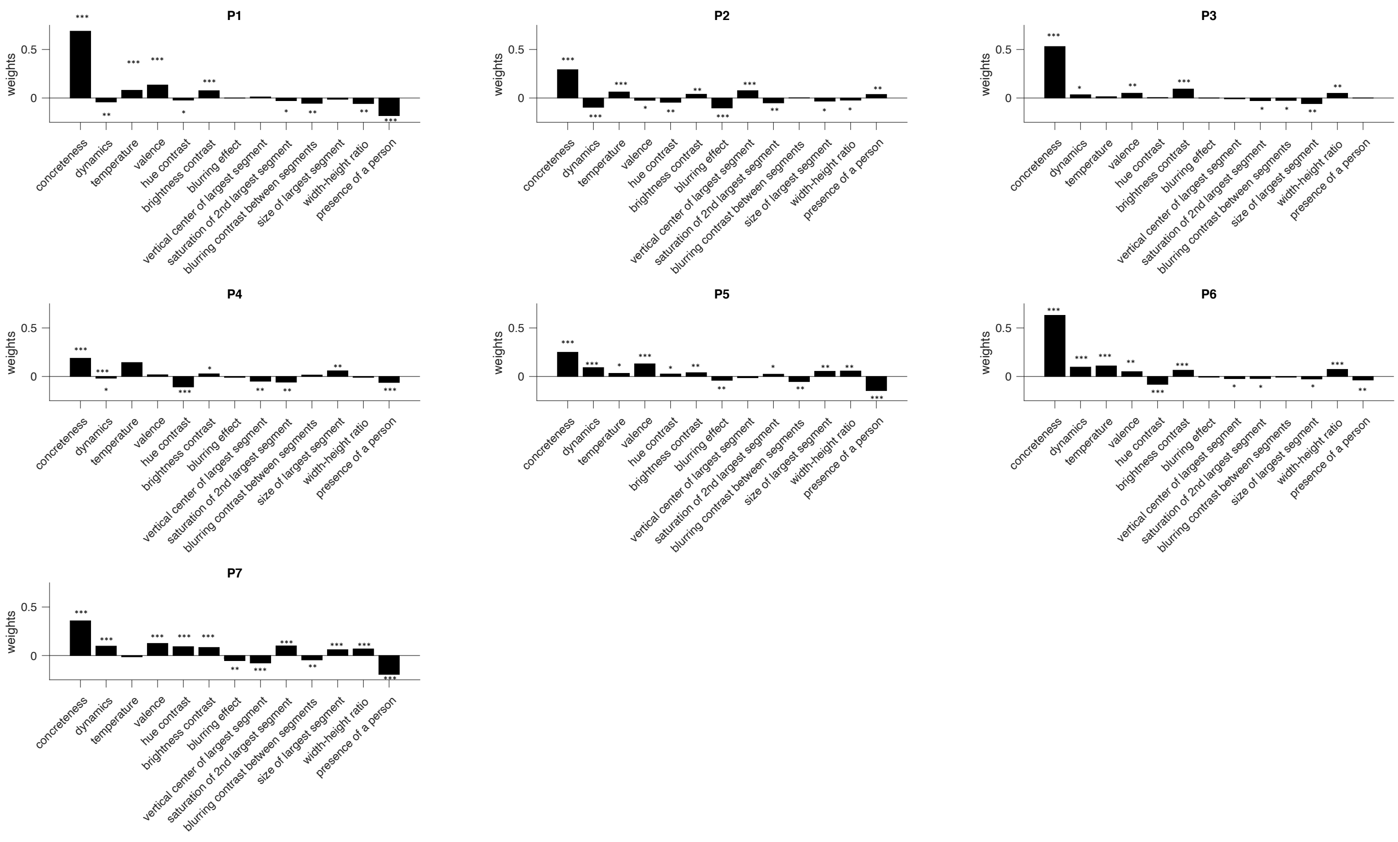
[1]Division of Humanities and Social Sciences, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125

*corresponding author: kiigaya@caltech.edu

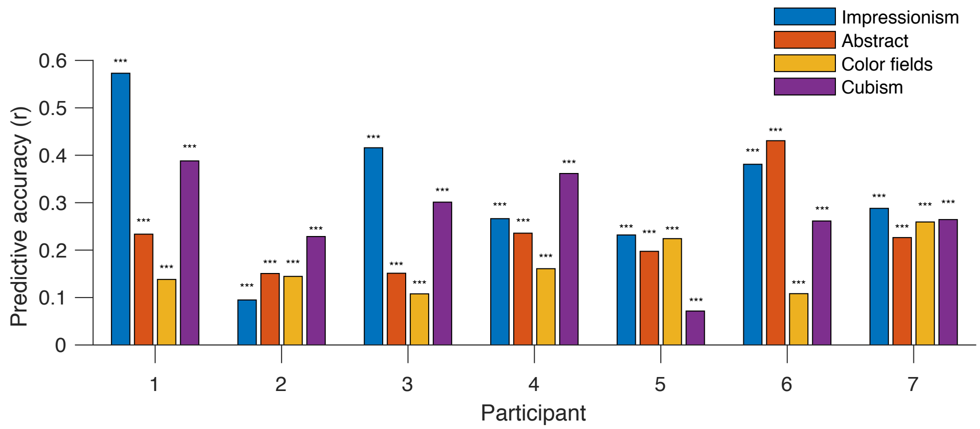
Supplementary figures



Supplementary Figure 1: The predictive accuracy of subjective ratings of art in the in-lab participants. Blue: within-participant prediction using our computational model. Red: across-participant prediction using the average rating of each stimulus.

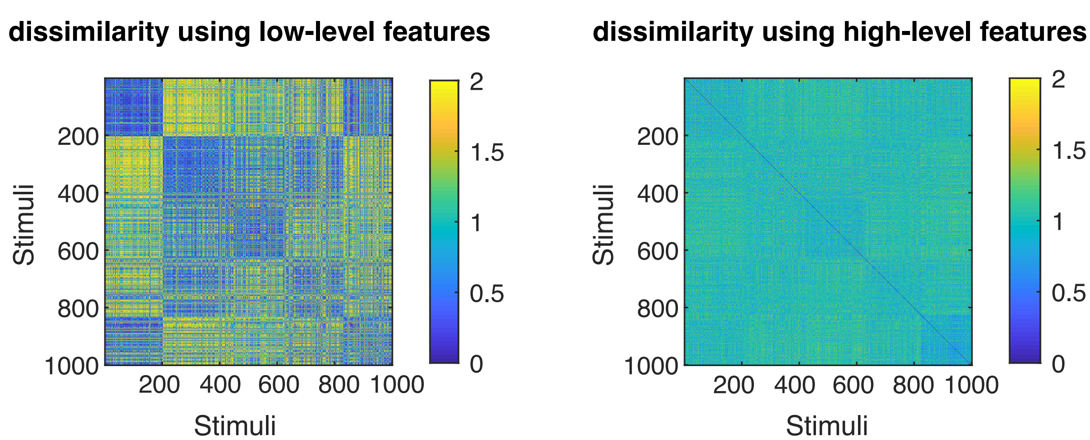


Supplementary Figure 2: The estimated feature weights of in-lab participants. The significance was estimated against the null distribution of weighs constructed by model fittings to permuted data. One star, two starts, three stars, indicates $p < 0.05$, $p < 0.01$, $p < 0.001$, respectively.

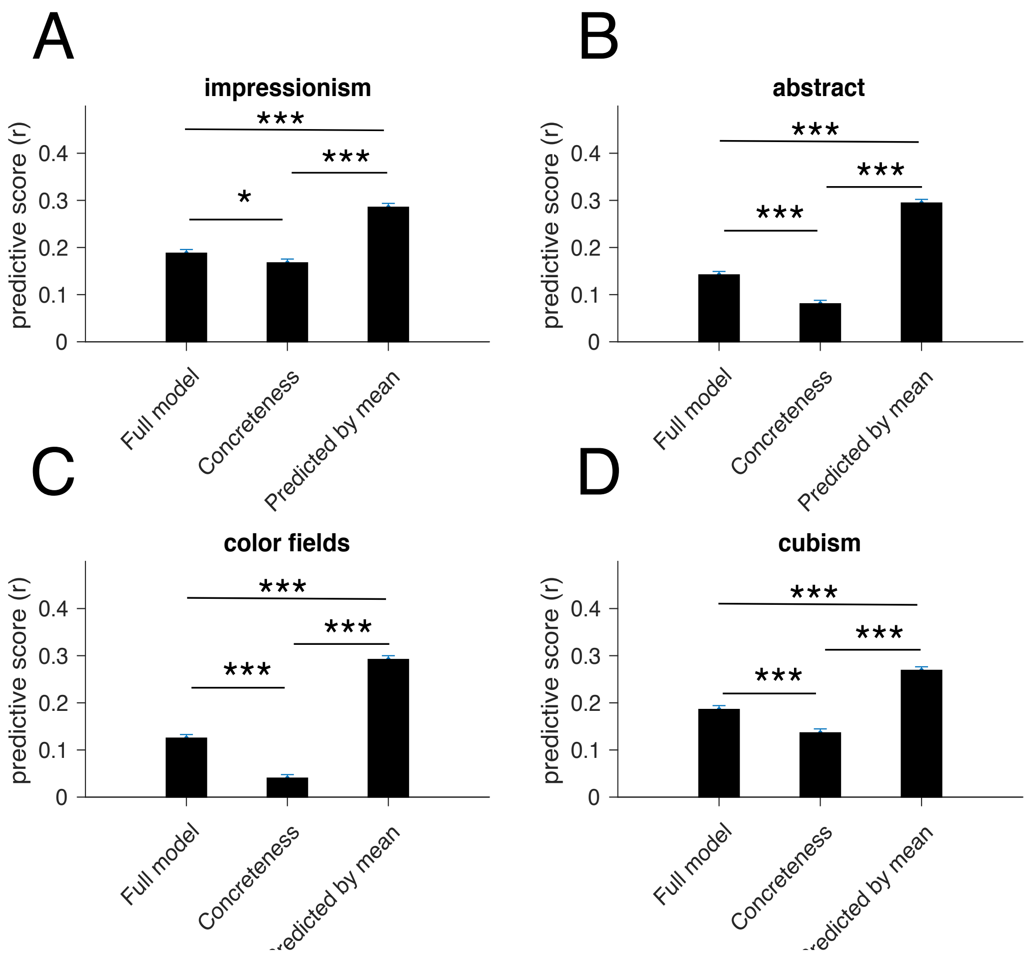


Supplementary Figure 3: Predictive accuracy of the LFS model within different art genres. The model was trained on all images using 20 fold cross validation in each participant. Predictions for images in each art genre were compared with the actual data. The predictive accuracy was measured by Pearson correlation. This figure shows that our overall predictive accuracy is not merely an artifact of the fact that people like different genres differently, i.e. that the LFS model is sensitive only to differences between images as a result of genre and that this alone enables it to have success. Here, even within specific genres, the model can still succeed in predicting liking ratings just as it can across genres. Note that correlation values are smaller than the overall value presented in Figure 1. This is because between-genre correlation is indeed present in Figure 1.

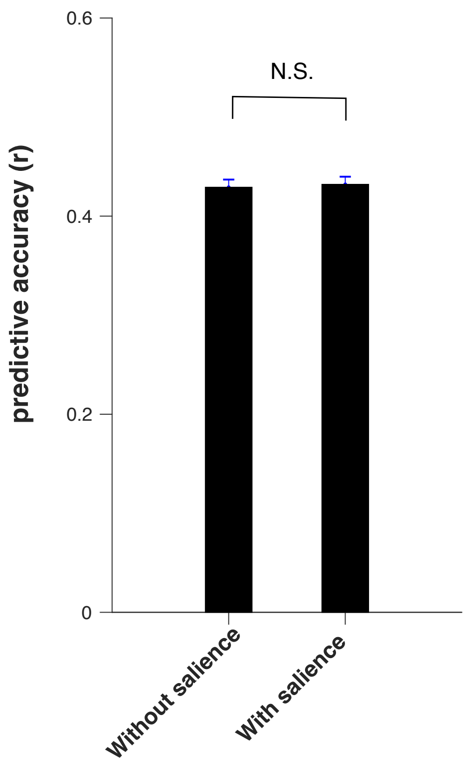
.



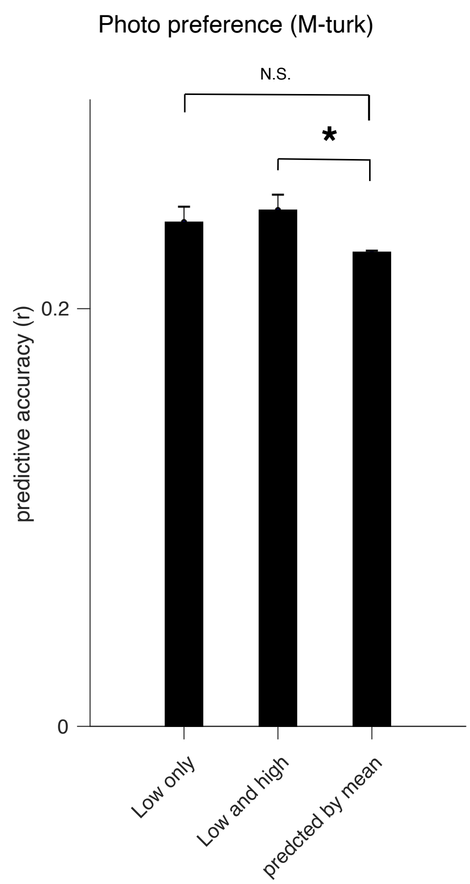
Supplementary Figure 4: Representation dissimilarity matrix using low-level and high-level features. Image index; 1-204: Impressionism. 205-417: Abstract art. 418-621: Color fields. 622-826: Cubism. 827-1000: Pictures from the stimulus set of Vaidya et al. (2017) .



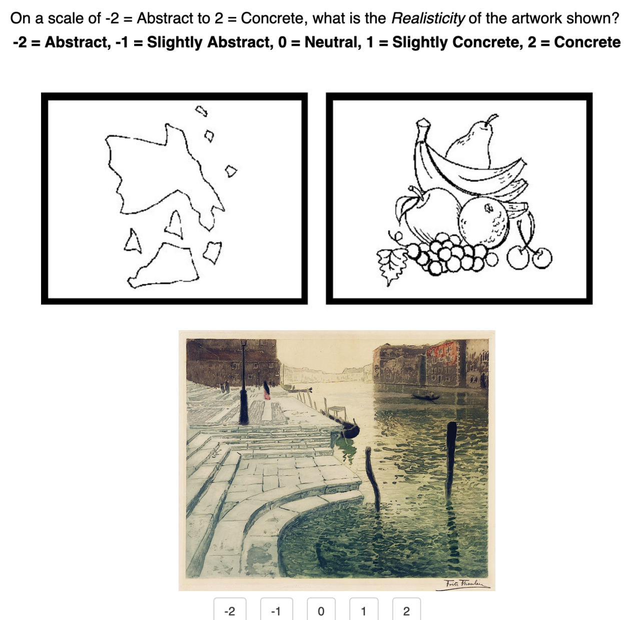
Supplementary Figure 5: The model's predictive accuracy when using full features or the concreteness feature alone, tested in the large scale online dataset. The full model significantly outperforms the model with concreteness feature alone, but shows room to improve when compared against the performance of an average rating model. The error bars indicate the mean and SEM.



Supplementary Figure 6: The inclusion of salience-weighted features does not improve our model's predictive accuracy in M-Turk participants. The error bars indicate SEM. N.S. indicates not significant in permutation test.



Supplementary Figure 7: The predictive accuracy of model on photograph ratings. Left: the original model with low-level features. Middle: the model with low-level and high-level features, where the binary high level features are approximated by a nonlinear support vector machine trained on visual art set using low-level features. Right: correlations with the average ratings for each image. The one star indicates $p < 0.05$ in permutation test across participants. The error bars indicate SEM.



Supplementary Figure 8: An example trial of feature annotation. Annotators were asked to evaluate high-level feature values (from -2 to 2).Frits Thaulow-Marmortrappen credit: ART Collection / Alamy Stock Photo.