Quantifying how humans trade off color and material in object identification

Ana Radonjić, Nicolas P. Cottaris and David H. Brainard
Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA.

Abstract

How do different object properties combine for the purposes of object identification? We developed a paradigm that allows us measure the degree to which human observers rely on one object property (e.g., color) vs. another (e.g., material) when they make forced-choice similarity judgments. On each trial of our experiment, observers viewed a target object paired with two test objects: a material match, that differed from the target only in color (along a green-blue axis) and a color match, that differed from the target only in material (along a glossy-matte axis). Across trials, the target was paired with different combinations of material-match and color-match tests and observers selected the test that appeared more similar to the target. To analyze observer responses, we developed a model (a two-dimensional generalization of the maximum-likelihood difference scaling method) that allows us to recover (1) the color-material weight, reflecting the relative importance of color vs. material in object identification and (2) the underlying positions of the material-match and color-match tests in a perceptual color-material space. Our results reveal large individual differences in the relative weighting of color vs. material.

Introduction

Object properties — such as color, material, texture or shape — all contribute to object identification and guide our interaction with objects in daily life. For example, color helps us select which tomato from the garden is ripe enough to eat and knowing whether a cup is made out of porcelain or plastic can help us decide how to handle it.

The visual processing underlying the extraction of each of these object properties from the retinal image has been extensively studied [1-5]. Little is known, however, about how these different properties combine to help us identify objects. To investigate this we developed a paradigm that allows us to measure the relative contribution of one property (e.g., object color) relative to another (e.g., object material) in identification.

In our method, we use an object selection task to measure how perceived object color and material trade-off in identification. The task is a generalization of the forced-choice color selection task we developed previously to study the stability of perceived object color across changes in illumination [6]. In the task, the observers are shown three objects — a target and two tests, all under equal illumination — and asked to select the "*test that is most similar to the target*". On each trial, the tests differ relative to the target in either color or material, to varying degrees. The criterion on which the observers are asked to make the similarity judgments is intentionally left unspecified.

To analyze the selection data we developed a model of selection behavior, which allows us to infer parameters that describe the perceptual representation of the stimuli and the relative weight observers place on perceived color versus perceived material. More specifically, the model parameters include: (1) the positions of the target and tests objects in a two-dimensional perceptual space (where the dimensions represent color and material) and (2) a color-material weight that determines how the overall perceptual distance between two stimuli depends on distances on each of the two-dimensions.

Methods

Apparatus

Stimuli were presented on a calibrated 27-in. NEC MultiSync PA241W LCD color monitor driven at a pixel resolution of 1920 x 1080 and at a refresh rate of 60 Hz , with eight-bit resolution for each RGB channel via an NVIDIA GeForce GTX 780M video card. The subject's head position was stabilized using a chin rest. The distance between the subject's eye and the center of the screen was 70 cm. The host computer was an Intel Core i7 Apple Macintosh. The experimental programs were written in Matlab and relied on routines from Psychtoolbox [7, 8] http://psychtoolbox.org and mgl (http://justingardner.net/doku.php/mgl).

Object identification task

On each experimental trial, three identical rendered scenes were displayed on the monitor, each containing a blob-shaped object (Figure 1). The object in the center scene was the target object, while the objects in the left and right scene were the test objects. The observers' task was to select the test object that was most similar to the target. They were instructed to use a mouse to "click" on the test object of their choice. A black dot then briefly flashed above the selected test to signal the response has been recorded. The screen then turned black for 1 second, after which the next trial started. Instructions verbatim are available in the online supplement (see below).

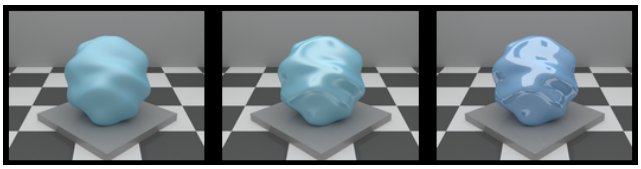

Figure 1. Example trial from the object identification task. *Observers select which of the two test objects (left or right) is most similar to the target object (center). Across trials, the degree of similarity of test objects relative to the target varies in both color (the material match, on the right, is bluer) and material (the color match, on the left, is more matte). The stimuli were presented against a black background (the background covered the entire display and is not shown here in its entirety).*

The experiment consisted of three different types of trials: (1) color-material-vary trials, (2) color-vary trials and (3) material-vary trails. On color-material-vary trials, one of the tests was always a *material match* for the target while the other test was a *color match* for the target. There were total of 48 color-material-

vary trials. The trail where the target is paired with two tests identical to itself was not included, as the performance for such a trial would be equal to chance, up to measurement variability. On color-vary trials, both tests were material matches and differed from the target only in color, each to a different degree (7 levels: from -3 to +3, including a 0 difference step). On material-vary trials, both tests were color matches and differed from target only in material, each to a different degree (also 7 levels). There were total of 21 material-vary and 21 color-vary trials (all pairwise comparisons of stimuli at 7 difference levels).

Stimuli

The stimulus scene was modeled in Blender, an open-source 3D modeling package (https://www.blender.org/). The scene consisted of room, whose walls were mid-gray. The floor was covered with a checkered pattern consisting of dark-gray and light-gray tiles. A blob-shaped object, resting on a square gray pedestal, was placed in the center of the room. The scene was illuminated by an area light that covered the entire ceiling. The illumination spectrum was a CIE daylight of 6500 K correlated-color temperature (D65). We generated a blob-shaped object from an icosahedral mesh that approximates a sphere. We subdivided each side of the mesh into 64 facets and then added a sinusoidal perturbation separately to the x, y and z coordinates of each facet vertex (with one sinusoid each for x, y, and z coordinates).

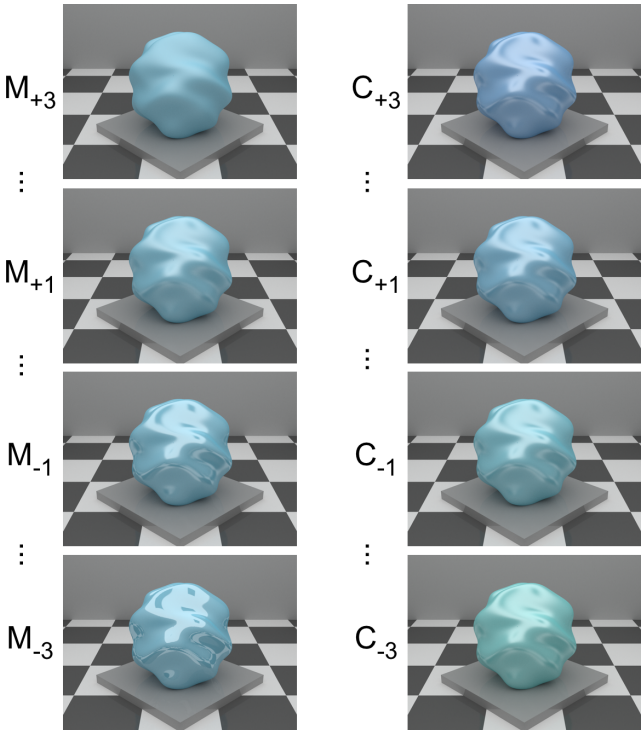

Figure 2. Example test stimuli. *Left column shows color matches (the tests are equal to the target in spectral reflectance but vary in gloss level). Right column shows material matches (the tests are equal to the target in material, but vary in spectral reflectance). Labels indicate the nominal degree of material/color difference relative to the target. Both columns show only 4 out of the 6 tests used in the experiment.*

We rendered total of 13 stimulus scenes. Scenes were identical in all respects, except for the target object characteristics. There were 1 scene with the target object, 6 with color-vary material matches and 6 with material-vary color matches. The color variation of the test objects was achieved by varying their assigned diffuse spectral reflectance component, by combining two different reflectance samples (a greenish one and a bluish one) in different proportions. For example, the target reflectance was 60% bluish and 40% greenish. For the remaining material matches diffuse reflectance mixtures had the following proportion of bluish reflectance sample (relative to the greenish one): 0.5 ($\Delta C = -3$), 0.53 ($\Delta C = -2$), 0.57 ($\Delta C = -1$), 0.63 ($\Delta C = +1$), 0.67 ($\Delta C = +2$), 0.70 ($\Delta C = +3$). The specular reflectance component for the material match test objects was fixed across all stimulus scenes (0.30 across all wavelengths).

The variation in gloss level of the test objects was achieved by varying the α_U and α_V parameters of their surface reflectance specification (in Mitsuba renderer notation, see below). These parameters control the anisotropic roughness of the material along the tangent and bitangent directions of reflected light [9]. Low α_U and α_V levels correspond to a glossy-appearing surface whose microstructure has small imperfections, while high levels correspond to a matte-appearing surface with rough microstructure. We set both α_U and α_V for the target object to 0.10. The levels for the remaining color matches were as follows: 0.007 ($\Delta M = -3$), 0.02 ($\Delta M = -2$), 0.05 ($\Delta M = -1$), 0.15 ($\Delta M = +1$), 0.20 ($\Delta M = +2$), 0.40 ($\Delta M = +3$).

We determined the spacing of tests in color and glossiness by eye, aiming to have adjacent tests along each dimension distinguishable from one another and the spacing between them roughly perceptually uniform. In the color domain, estimated differences between two nearest tests under our experimental illumination were 5.33 CIELAB ΔE units, on average (with standard deviation of 0.15 ΔE).

The stimuli were rendered in Mitsuba, a physically-based rendering package, (http:// www.mitsuba-renderer.org/), using a bidirectional path tracer integrator and low discrepancy sampler (sample count: 1024). RenderToolbox3 [10] routines were used to facilitate the rendering and to assign surface reflectance functions and illumination spectra to the elements in the scene. Each rendered stimulus scene led to a 31-plane hyperspectral image (960 x 720 pixels per plane). The hyperspectral images were converted into a three-plane LMS image by computing the excitations that would be produced in the human L-, M-, and S-cones at each pixel (using Stockman–Sharpe cone fundamentals [11, 12]). Then, the LMS images were converted into RGB stimulus images for presentation, based on display calibration measurements and using standard methods [13]. Calibration measurements were made using a PhotoResearch PR-670 spectral radiometer and included the characterization of spectral power distribution of the display primaries as well as the display gamma function for each channel.

At the 70 cm distance from the screen, each stimulus image subtended 14.7° x 11° degrees of visual angle (18 x 13.5 cm). The size of each test object was approximately 6.7° x 6.5° (8.2 x 7.9 cm).

Observers

Five observers (University of Pennsylvania undergraduates) participated in the experiment (4 female, 1 male; age: 18-19). They all had normal or corrected-to-normal visual acuity (20/25 or better in both eyes, as assessed via a Snellen chart) and normal color

vision (0 plates incorrect on Ishihara color plates) [14]. Observers received course credit for their participation.

Experimental Procedures

Each observer completed 25 blocks of trials, over 4 sessions (6-7 blocks per session). Each session lasted approximately an hour and was run on a different day. Each experimental block consisted of 76 trials (48 color-material-vary trials, 21 color-vary and 21 material-vary trials).

At the beginning of an experiment observers completed a brief training block, consisting of four trials. In the training trials, one of the test objects was always identical to the target while the other test was either the most-different color match (ΔM = +/-3) or the most-different material match (ΔC = +/-3).

Preregistration and supplementary materials

A document describing the experimental design and the data analysis plan for this study was preregistered before the start of data collection and is publically available at https://osf.io/jtqs2/. The model we present was developed after data collection was completed and is not described in detail in the preregistration.

Supplementary materials available online (URL: http://color.psych.upenn.edu/supplements/colormaterialEI2018) include rendering materials (surface reflectance functions and illumination spectra, the model of scene geometry, conditions and mappings files used to assign properties to each element in the scene) as well as the selection data and visualization of model fits (corresponding to Figures 3 and 4 below) for each observer.

Results

Aggregating the data

Figure 3 shows the results for one observer, for the color-material-vary trials. Each point shows data for one color match/material match pair, and plots the proportion of trials on which the color match was chosen. The symbol color indicates the levels of material difference for the color match, while abscissa gives the color difference of the material match. The smooth lines show the fit of our model, which we describe below. Note that although only a subset of data is shown in the figure, the full data set was used in fitting the model. Corresponding figures for all observers are provided in the online supplement.

To understand Figure 3, consider first the black symbols, which indicate observer choices when the color match identical to the target (ΔM = 0) is paired with material matches that vary in color (x-axis). As the material match deviates from the target (points away from ΔC = 0 on the x-axis), the observer consistently chooses the color match, as one would expect. Next consider the red symbols, which represent a color match that differs considerably in material from the target (ΔM = +/-3). When the material match is identical to the target (ΔC = 0), then the observer always selects the material match (ordinate of 0), again as one would expect. As the color difference of the material match increases, however, the proportion of trials on which the color match is chosen also increases, as now the observer must select between two tests that each differ from the target, but in different ways. The rate at which the selections in this case transition from material match to color match choices is related to how the observer trades off color and material differences in selection. Finally, for the small and medium material difference steps of the color match (ΔM = +/-1 in blue; ΔM = +/-2 in green) the data fall in between the two extremes. This shows that the smaller the difference in material of the color match, the faster the observer transitions between selecting the material match to selecting the color match. Again, this makes intuitive sense.

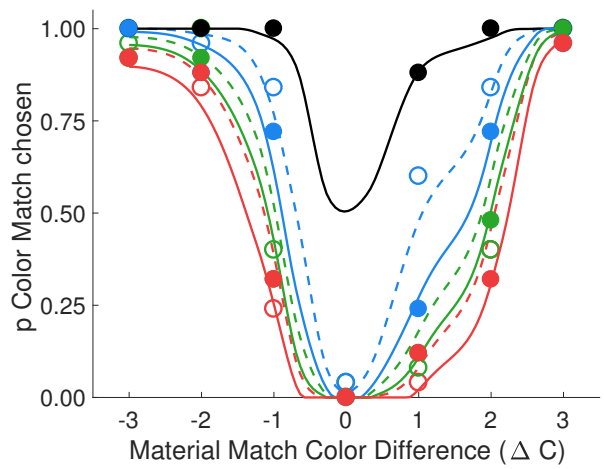

Figure 3. Data and model fit for one observer (zhr). *The proportion of trials on which the color match was chosen is plotted against the size of the deviation in color between the material match and the target object. Each set of points (black, blue, green, red) shows data for a different magnitude of material deviation of the color match from the target (zero deviation is plotted in black, small in blue, medium in green and large in red). Open circles show negative and full circles show positive differences. The smooth lines show the model fit to the data. Model prediction line colors correspond to the symbol colors used to plot the same material difference level of the color match. Dashed lines show fit to open circles (negative difference steps); solid lines show fit to full circles (positive difference steps). The black curve is, sensibly, forced by the structure of the model to pass through ordinate 0.5 at abscissa 0, which corresponds to the hypothetical case of two tests are identical to the target in both color and material* (ΔM = 0 and ΔC = 0).

Modeling the data

Although Figure 3 provides a qualitative sense of how color and material trade off in selection, quantitative statements require a model. To this end, we extended our earlier work on modeling of object selection based on color [6, 15, see also 16] to handle the case where both color and material vary. The model fits our full data set in terms of small number of interpretable parameters that efficiently characterize the underlying perceptual representation of the stimuli as well as the relative weighting of color and material used in selection. The model assumes that on each trial, each stimulus — the target and test objects — is represented as a point in a two-dimensional perceptual space, where the dimensions represent variation in color and material respectively. Because perceptual representations are noisy [17], the model characterizes each stimulus as a noisy distribution centered around its mean position in the perceptual space. To model what happens on a trial, the trial-specific positions of the target and two tests are obtained as draws from their underlying distributions, and the observer's choice is determined as a function of these trial-specific positions. More specifically, in the model observer chooses the test whose position is closest to the target on that trial, with distance obtained using a Euclidean metric that differentially weights differences in the color and material dimensions. In other words, we compute the distance from the target to the first test ($d_{T\text{-}T1}$) as:

$$(1)\quad d_{T-T_1} = \sqrt{(w(\Delta C_{T-T_1}))^2 + ((1-w)(\Delta M_{T-T_1}))^2}$$

where ΔC denotes color difference between the target T and the first test T_1, weighted by a color-material weight w and ΔM denotes a material difference between target and the first test, weighted by $1\text{-}w$. Analogously, the distance from the target to the second test T_2 ($d_{T\text{-}T2}$) is computed as:

$$(2)\quad d_{T-T_2} = \sqrt{(w(\Delta C_{T-T_2}))^2 + ((1-w)(\Delta M_{T-T_2}))^2}$$

On a given trial, the observer selects the first test if $d_{T\text{-}T1} < d_{T\text{-}T2}$ and the second test otherwise.

In fitting the model, we lock the mean position of the target at [0,0] and set the Gaussian perceptual noise of the two dimensions (σ_C, σ_M) to be equal to 1. These choices determine the origin of the perceptual space and set the scale of the two axes to be equal when we perceptual distance is characterized in units proportional to perceptual precision. The model then has thirteen free parameters. One parameter (the weight w) describes the color-material trade-off. The other twelve free parameters describe the mapping from physical stimuli to the perceptual representation. That is, each parameter defines the position of one of the test stimuli in the two-dimensional color-material space (6 color matches and 6 material matches). For each of the 6 color matches, the free parameter controls only the position on the material dimension, while the position on the color dimension is set to be equal to that of the target (0). Similarly, for each of the 6 material matches, the free parameter controls only the position on the color dimension, while the position on the material dimension is equal to that of the target (0). These parameters accommodate the possibility that in the perceptual space, the stimuli are not distributed uniformly along either axis.

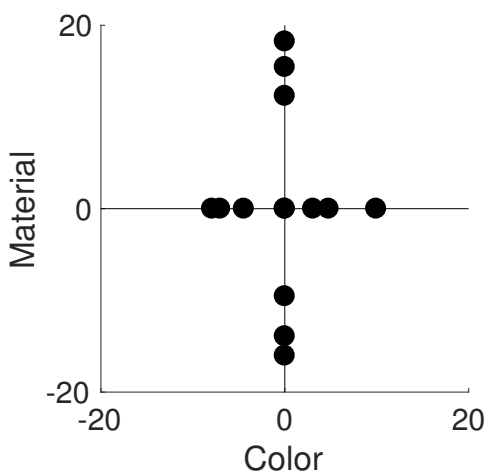

Figure 4. Inferred stimulus positions in the perceptual color-material space for one observer (zhr). *Circle in the center indicates the target position, which was fixed in our model. Circles along the x-axis indicate the position on the color dimension for each material match (from ΔC -3 on the left to ΔC +3 on the right). Circles along the y-axis indicate the position on the material dimension for each color match (from ΔM -3 on the bottom to ΔM +3 on the top). The inferred color-material weight for this observer was 0.76.*

For each observer, we fit the model's parameters to the data via numerical search, using a maximum likelihood criterion. Figure 4 shows the stimulus positions obtained from the model fit for observer *zhr*, whose data are shown in Figure 3, along with the predictions from the model (smooth curves in figure). The inferred color-material weight for this observer is equal to 0.76. Clearly, for this observer the model accounts well for the data. Fits for the remaining four observers are shown in the online supplement, together with the corresponding inferred stimulus positions.

We estimated the precision of the inferred color-material weights using a bootstrap procedure [18]. For each observer, we generated new selection data sets (78 trials x 25 blocks) by resampling responses (with replacement, from the original data set), for each trial type independently. We repeated this procedure until we obtained 150 sets of resampled data. We then searched for the best-fitting model parameters for each set. Figure 5 shows for all observers the mean inferred color-material weight, across the 150 repetitions of resampling of the original data (green symbols), with error-bars indicating +/-1 standard deviation of the 150 repetitions. Black symbols show the color-material weight inferred, by our model, from the actual set of observer's responses.

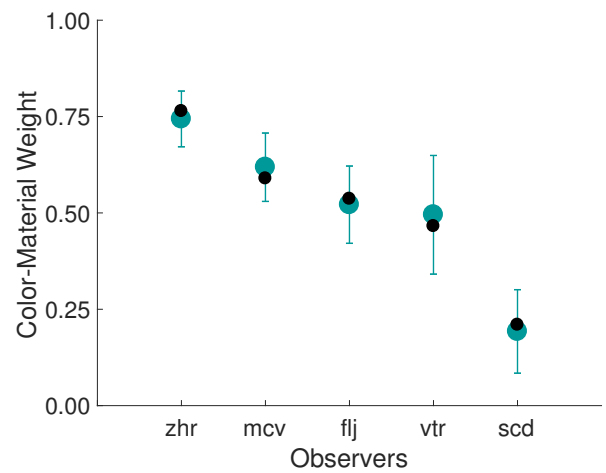

Figure 5. The relative importance of color and material as cues to object identity varies across observers. *The mean bootstrapped weight (across 150 repetitions) is shown for each of our five observers (green symbols). Larger weights indicate observers who rely more heavily on color relative to material. Error bars determined through bootstrapping show +/- 1 standard deviation. Black symbols indicate the color-material weight inferred from the actual selection data set for each observer.*

Figure 5 illustrates a key feature of our dataset: the color-material weight varies considerably across observers. For observer *zhr* the color-material weight is high, indicating that this observer relied predominantly on color when making similarity judgments. In contrast, for observer *scd* the color-material weight is low, indicating that selection is predominantly driven by object material. The weights for the remaining three observers are similar and indicate more balanced color-material weighting in selection. Although the error bars are larger for some subjects than others, those between *zhr* and *scd* do not overlap, suggesting that these differences in weight are statistically reliable.

Discussion

The work we describe provides the first quantification of color-material weighting for object identification and indicates that there are large individual differences in how these two object properties are used. A key innovation of the work is the development of a model that allows us to derive both the perceptual representation of the stimulus set and the color-material weight, so that our conclusions do not require a priori assumptions about the perceptual distances between the stimuli along the color and material axes. In recent related work, Saarela and Olkkonen showed that the visual system can combine information about object color and glossiness to improve performance in an object discrimination task [19]. This finding is consistent with our result showing that observers use both color and material in selection.

We investigated color-material trade off using a single type of object material change, a single object color range and a single object shape. It would be interesting to expand the range of stimulus variations studied. It would also be interesting to use dynamic stimuli and have observers compare target and test objects while they rotate, each with a different orientation phase. This would preclude direct comparisons of specific regions of the images (e.g., the specular highlight at one image location) and encourage observers to make their similarity judgments based on global object appearance.

An intriguing question for future work is whether, for a particular subject, the relative weighting of color and material varies in response to variation in their relative reliability. Color might be less reliable when the target and tests are viewed across a change in the spectral properties of the illumination, while material might be less reliable when there is a change in the geometric properties of the illumination under which the object is viewed.

There are some limitations of the current modeling. First, our model is based on assumptions that may be overly strong. For example, we assume that the material matches have their mean positions aligned precisely along the perceptual color axis and that the color matches lie precisely along the perceptual material axis. Previous work has shown that the perceptual representation of physical surface glossiness varies with changes in the physical roughness of an object surface, and that perceived roughness varies with changes in physical glossiness [20], so we are alert to this possibility for color and material [21]. In addition, we neglect noise added to the distance comparison and place all of the model variability in the perceptual stimulus representation. Indeed, there are maximum likelihood approaches to multidimensional scaling that allow for decision noise, but not for perceptual noise [22, 23]. Finally, we have so far restricted attention to a Euclidean distance metric, but it may also be interesting to consider other metrics (e.g., “City-Block” [24]).

Relaxing model assumptions is not trivial, however. As we have proceeded, we have found that adding model parameters to account for realistic factors increases the amount of data required to estimate model parameters with sufficient precision, as parameters can be co-varied with only a small effect on the model’s predictions. We have begun to approach this by including tests that differ from the target in both color and material, and by using an adaptive psychophysical procedure [25] to choose the test pairs to be presented on each trial. In the end, however, the complexity of the models we can consider and the precision to which parameters can be estimated may be limited by considerations of parameter identifiability. This in turn means that it is important to frame our conclusions within the context of the specific modeling assumptions we make.

Acknowledgements. This work has been supported by NIH RO1 EY10016, NIH Core Grant P30 EY001583 and the Simons Foundation Collaboration on the Global Brain Grant (324759).

Author Biographies

Ana Radonjić is a Research Associate at the Psychology Department at University of Pennsylvania. She received a B.A. in Psychology from University of Belgrade, Serbia (2001), and an M.A. and a Ph.D. in Psychology from Rutgers University (2009). She studies mechanisms underlying human perception of surface color and material using naturalistic stimuli and tasks.

Nicolas P. Cottaris is a Senior Scientific Application Developer at the University of Pennsylvania. He received a B.S. in Computer Science from the University of Crete, Greece (1989), an M.S. in Biomedical Engineering from Rutgers University (1992) and a Ph.D. in Vision Science from UC Berkeley (2001). His current research focuses on computational modeling of the early stages of human visual processing.

David H. Brainard is the RRL Professor of Psychology at the University of Pennsylvania. He received an AB in Physics from Harvard University (1982) and an MS (Electrical Engineering) and PhD (Psychology) from Stanford University (1989). His research focuses on color vision and color image processing. He is a fellow of the Optical Society, ARVO and the Association for Psychological Science, a director of Penn's Vision Research Center, and a co-director of Penn's Computational Neuroscience Initiative.

References

[1] D. H. Brainard and A. Radonjić, "Color constancy.," in *The New Visual Neurosciences*, Werner J. S. and L. M. Chalupa, Eds., ed Cambridge, MA: MIT Press, 2014, pp. 545-556.

[2] D. H. Foster, "Color constancy," *Vision Research,* vol. 51, pp. 674-700, 2011.

[3] A. C. Chadwick and R. W. Kentridge, "The perception of gloss: A review," *Vision Research,* vol. 109, pp. 221–235, 2015.

[4] R. W. Fleming, "Visual perception of materials and their properties," *Vision Research,* vol. 94, pp. 62-75, 2014.

[5] J. Koenderink and A. J. van Doorn, "Shape and Shading," in *The Visual Neurosciences.* vol. 2, L. Chalupa and J. S. Werner., Eds., ed Cambridge, MA: MIT Press, 2004, pp. 1090-1105.

[6] A. Radonjić, N. P. Cottaris, and D. H. Brainard, "Color constancy supports cross-illumination color selection," *Journal of Vision,* vol. 15, pp. 1-19, 2015.

[7] D. H. Brainard, "The Psychophysics Toolbox," *Spatial Vision,* vol. 10, pp. 433-436, 1997.

[8] D. G. Pelli, "The VideoToolbox software for visual psychophysics: transforming numbers into movies," *Spatial Vision,* vol. 10, pp. 437-442, 1997.

[9] J. Wenzel, "Mitsuba Documentation (Version 0.5.0)," 2014.

[10] B. S. Heasly, N. P. Cottaris, D. P. Lichtman, B. Xiao, and D. H. Brainard, "RenderToolbox3: MATLAB tools that facilitate physically based stimulus rendering for vision research," *Journal of Vision,* vol. 14, 2014.

[11] A. Stockman and L. T. Sharpe, "Spectral sensitivities of the middle- and long-wavelength sensitive cones derived from measurements in observers of known genotype," *Vision Research,* vol. 40, pp. 1711-1737, 2000.

[12] CIE, *Fundamental chromaticity diagram with physiological axes – Parts 1 and 2. Technical Report 170-1*. Vienna: Central Bureau of the Commission Internationale de l' Éclairage, 2007.

[13] D. H. Brainard, D. G. Pelli, and T. Robson, "Display characterization," in *Encyclopedia of Imaging Science and Technology*, J. P. Hornak, Ed., ed New York, NY: Wiley, 2002, pp. 172-188.

[14] S. Ishihara, *Tests for Colour-Blindness*. Tokyo: Kanehara Shuppen Company, Ltd., 1977.

[15] A. Radonjić, N. P. Cottaris, and D. H. Brainard, "Color Constancy in a naturalistic, goal-directed task," *Journal of Vision,* vol. 15, pp. 1-21, 2015.

[16] L. T. Maloney and J. N. Yang, "Maximum likelihood difference scaling," *Journal of Vision,* vol. 3, pp. 573-585, 2003.

[17] C. Blakemore, *Vision: Coding and Efficiency*. Cambridge: Cambridge University Press., 1990.

[18] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap.* New York: Chapman and Hall/CRC 1993.

[19] T. Saarela and M. Olkkonen, "Integration of color and gloss in surface material discrimination," *Journal of Vision,* vol. 17, pp. 229-229, 2017.

[20] Y. X. Ho, M. S. Landy, and L. T. Maloney, "Conjoint measurement of gloss and surface texture," *Psychol Sci,* vol. 19, pp. 196-204, Feb 2008.

[21] S. Hansmann-Roth and P. Mamassian, "Simultaneous gloss contrast: Conjoint measurements of lightness and gloss," *Journal of Vision,* vol. 16, pp. 943-943, 2016.

[22] J. O. Ramsay, "Maximum likelihood estimation in multidimensional scaling," *Psychometrika,* vol. 42, pp. 241-266, 1977.

[23] Y. Takane, "A maximum likelihood method for nonmetric multidimensional scaling: I. The case in which all empirical pairwise orderings are independent – theory and evaluations," *Japanese Psychological Research,* vol. 20, pp. 7-17, 1978.

[24] D. Algom and D. Fitousi, "Half a Century of Research on Garner Interference and the Separability–Integrality Distinction," *Psychological Bulletin,* vol. 142, pp. 1352–1383, 2016.

[25] A. B. Watson, "QUEST+: A general multidimensional Bayesian adaptive psychometric method," *Journal of Vision,* vol. 17, pp. 1–27, 2017.