Color spaces of color-normal and color-abnormal observers reconstructed from response times and dissimilarity ratings

GALINA V. PARAMEI and C. RICHARD CAVONIUS
Universität Dortmund, Dortmund, Germany

With multidimensional scaling analysis, color spaces were reconstructed from reaction times (RTs) required to make *same-different* judgments of pairs of 15 equiluminant colors and from dissimilarity ratings between them. In addition to normal trichromats, observers with red-green color deficiency were tested. Two main purposes were served by this study: (1) to compare spatial representations of colors derived from discriminative RTs with those derived from dissimilarity measures and (2) to examine whether the task may selectively affect the dimension reflecting, in color-abnormal spaces, the deficient red-green mechanism. Contradicting our hypothesis of lower dimensionality of RT spaces, as compared with rating spaces, no consistent differences in solution dimensionality were found. However, configurations derived from the two measures diverged. The rating procedure yielded the most logical results for recovering color space. The RT configuration revealed contraction in the tritanopic direction, indicating longer color processing when the short-wavelength mechanism is involved, and in addition, for color-abnormal observers, clustering in the protanopic and deuteranopic directions, indicating even longer processing by the deficient red-green mechanism. This finding implies that RTs are suitable for detecting temporal differences in color processing but, for that very same reason, rather ill-suited for reconstructing color spaces.

A long-standing problem in color perception is the psychological structure onto which color dimensions may be mapped. Among the current approaches to uncovering such a mapping is the study of color differences among a set of stimuli embedded in a space with an appropriate number of dimensions, using multidimensional scaling (MDS) analysis (Young & Hamer, 1987). Color spaces reconstructed in this way are uniform, in the sense that equal distances between any two points are thought to represent equal subjective differences between the corresponding colors. Despite a common methodology, different studies have obtained spatial solutions that diverge in dimensionality, with significant implications for the possible neural processes underlying color perception—in particular, the postreceptoral processing represented by the corresponding dimensions. From a formal point of view, divergent

solutions are admissible, for in an MDS analysis, the decision as to how many dimensions are appropriate is based on their interpretability and on the measure that evaluates departure from a perfect monotonic relation between differences and distances—namely, *stress* (Kruskal, 1964). This measure can conceivably depend on the particular measure of color difference employed in different tasks.

This article presents data for the comparison of color spaces reconstructed from two measures of color similarity, using the same stimuli and the same observers. The impetus for the present study was a discrepancy in the dimensional solutions obtained in two laboratories (that of Cavonius and Mollon and that of Sokolov and collaborators) with comparable sets of stimuli, viewing conditions, and MDS algorithms and with observers with color-normal and color-abnormal vision. The main difference between the two studies concerned the experimental task and the similarity measure: In the former study, *discriminative choice reaction times* (RTs) obtained in *same-different* judgments served as measures of the *discriminability* of colors, whereas in the latter, *ratings* reflected *dissimilarities* rooted in generalization (Shepard, 1986). Although both tasks involved partitioning of stimuli into *same* and *different* categories, Blough (1988) argues that it is not obvious that a decision process proposed by Shepard for generalization would apply to discrimination. “Instead of being determined by the chance that a stimulus is in a consequential region, RT would seem to be determined by how fast any difference at all can be detected” (Blough, 1988, p. 68).

This research was supported by DFG Grants 436 RUS 17/202/93S to G.V.P. and Ca 126/3-1 to C.R.C. The authors very much appreciate the cooperation of D. Bimler, who kindly extended the analysis of our data, using MDS algorithms elaborated by him. We are grateful to E. N. Sokolov and T. Indow for useful discussions and to W. Backhaus and the late A. P. Petrov for comments on an earlier version of the article. The authors' thanks are extended to J. D. Mollon, J. Theeuwes, and two anonymous reviewers for valuable criticisms of the first version of this article. The assistance of P. Dillmann, P. Schwarzenau, C. Westedt, and U. Lobisch is highly appreciated. Correspondence concerning this article should be addressed to C. R. Cavonius, Universität Dortmund, Institut für Arbeitsphysiologie, Ardeystrasse 67, D-44139 Dortmund, Germany (e-mail: cavonius@arb-phys.uni-dortmund.de).

—Accepted by previous editor, Myron L. Braunstein

The purpose of the present study was to find out whether the two methods yield different color spaces (number of dimensions, their interpretation, and their goodness-of-fit). The rationale was that the time allotted to the subject to respond, as well as the decision process, are additional factors that may influence the processing of stimulus dimensions. Besides normal trichromatic vision, congenital red-green deficiencies were examined, in order to test the interaction of the particular *color-processing capacities* with different methodological conditions.

Since the earliest MDS analyses by Shepard (1962) and Helm (1964), color spaces for normal trichromats derived from the similarity among hues of constant brightness and saturation (Ekman, 1954) are considered to be sufficiently described as a circular configuration in two dimensions, which Helm identified as blue-yellow and red-green, in accordance with the opponent-process theory of Hurvich and Jameson. Two-dimensional solutions were also preferred by Cavonius and Mollon (1984), with equiluminant lights, using discriminative RTs. One dimension was interpreted as red-green, and the other as violet-yellow (tritanopic), the latter corresponding to one of the cardinal directions of Krauskopf, Williams, and Heeley (1982). This seemingly minor difference in labeling the dimension having purplish-blue and yellow as extremes has implications for the second-stage processing of color, since blue-yellow opponency implies that the red-green system is at equilibrium, whereas a tritanopic (violet-yellow) dimension implies that short-wavelength cones are modulated under a constant ratio of long- and middle-wavelength cone excitation, and not a red-green equilibrium.

Using a quite different method of color difference evaluation—rating of the perceived distinctness of a border between lights of constant luminance—only one chromatic dimension, red-green, emerged, accompanied by a second dimension tentatively related to saturation (Tansley & Boynton, 1976). Indiscriminability of colors along the tritanopic dimension is thought to result from the poor spatial acuity of short-wavelength cones, which contribute little to the perception of contour.

Unlike the previous findings, Sokolov and Izmailov (1983), who had their subjects rate dissimilarities between equiluminant lights, concluded that a three-dimensional solution was appropriate on the basis of the number of large characteristic roots. With a conventional circular configuration, the first two dimensions were interpreted in terms of red-green and blue-yellow systems, associated with the two opponent processes in the Hurvich-Jameson theory. The loci of the points along the third dimension formed a surface with a constant positive curvature—that is, a hemisphere—so that the unipolar dimension was considered to characterize saturation of equally bright colors (although, in their stimulus set, equiluminant lights were used). Their three-dimensional solution found support in Judd (1966), who pointed out that although an overwhelming majority of studies represented equilight colors by a planar array, geometrical difficulties are encountered that indicate that color space is

curved, a conclusion that goes back to suggestions of Helmholtz and Schrödinger (see Shepard, 1980).

The color spaces for red-green-deficient observers that were obtained in Helm's (1964) study were also considered to be two-dimensional. The two axes of the deuteranomalous (DA) color spaces were interpreted as blue-yellow and red-green systems, in the order of their significance. In the case of a protanomal (PA), the first dimension was interpreted as blue-yellow, but the second was not identified. Helm noted, however, that the data of color-deficient observers were clearly more than two-dimensional, with obvious distortions imposed on the structure. In Helm and Tucker (1962), three dimensions were retained for further analysis. A marked contraction of distances in the red-green direction was observed in all color-deficient spaces. In addition, one can notice that, in more severe forms of color deficiency, the dimension that is interpretable as red-green appears to be third in the order of significance (see Helm & Tucker, 1962, Figure 3). An individual-differences MDS reanalysis of Helm's data for color-deficient observers, together with those for normal trichromats (NTs), resulted in two-dimensional solutions, with a major blue-yellow and a minor red-green dimension (Borg & Lingo, 1987; Chang & Carroll, 1980).

Spatial solutions for anomalous trichromats, reconstructed from RTs (Cavonius, Müller, & Mollon, 1990; Müller, Cavonius, & Mollon, 1991) are not quite consistent with the previous findings. A set of equiluminant stimuli that, for normal observers, yielded an approximately checkerboard pattern in two dimensions gave a degenerate solution for the PAs and the DAs: a horse-shoe, along which the stimuli were ordered according to their salience for the short-wavelength cones. The anomalous color spaces seem not to exhibit a second axis corresponding to the red-green system. To explain the horse-shoe configuration, the authors noted that unidimensional physical stimuli might correspond to a bipolar perceptual dimension of saturation.

Paramei, Izmailov, and Sokolov (1991) obtained color spaces for a protanope and an extreme DA, using dissimilarity ratings of equiluminant lights, and identified three dimensions: red-green, blue-yellow, and saturation—that is, analogues to those in the normal trichromatic space (although, again, it remains possible that the "saturation" dimension actually reflected brightness inequality among stimuli). The configurations were also found to lie on a hemispherical surface. The protanopic configuration was compressed along the red-green axis and stretched along the blue-yellow; in addition, the dimension interpreted as saturation had a higher significance than did the red-green. The deuteranomalous configuration differed, in that all the points were displaced toward the *white* pole along both chromatic dimensions. As has been shown elsewhere (Paramei, 1982), the degree of contraction of the configurations, for both protans and deutans, corresponded to the degree of color deficiency diagnosed by traditional methods.

A common feature of all spatial solutions for red-green-deficient observers is the perceptual salience of

Figure 1. The set of chromaticities plotted in the CIE x,y diagram. The coding of the colors is as follows: R, red; O, orange; Y, yellow; GY, green-yellow; G, green; GB, green-blue; B, blue; BP, blue-purple; P', purple ('); I, indigo; IP, indigo-purple; PI, purple-indigo; P'', purple ("); PR, purple-red; W, white. Solid lines indicate tritanopic confusion lines, dotted lines indicate protanopic confusion lines, and dot-dashed lines indicate deuteranopic confusion lines.

the dimension spanning yellow at one extreme and blue or violet at the other. The discrepancy across studies is concerned with the dimensionality of color-abnormal spaces: Those derived from discriminative RTs yield degenerate one-dimensional solutions, whereas spaces derived from (dis)similarity judgments are two- or even three-dimensional.

The review of the previous work led us to suggest that dimensionality differences of reconstructed color spaces may depend on the method and/or the measure of color differences used. In particular, we hypothesized that the speeded *same-different* procedure might result in lower dimensionality than does dissimilarity rating. In addition, in spaces for color-abnormal observers, the former procedure might be expected to selectively affect discrimination provided by the deficient red-green mechanism and to reduce the corresponding dimension.

METHOD

Observers

The observers were 3 NTs, including the authors, and 3 color-deficient subjects, 2 of whom, H.A.N. and J.O.N., were siblings. Phenotypes refer to the classification obtained with the help of the Ishihara (1991) plates and the Nagel anomaloscope (Model I, Schmidt & Haensch). Both siblings were diagnosed as PAs. The subject E.H.R., who accepts a match at both extremes of the Nagel

scale, with scale settings of the yellow light to match the green and the red primaries at 14.5 and 16, respectively, was diagnosed as deuteranope (D). The female subject, G.V.P., showed an enlarged range on the Nagel scale and misread three of the plates, the errors in which suggest a carrier of color deficiency (Jordan & Mollon, 1993).

All the observers participated in experiments with both *same-different* and dissimilarity judgments. They had preliminary training, to familiarize them with procedures and to improve their ability to quantify perceived color differences. The subject's head position was maintained with a chinrest; observation was binocular.

Stimuli and Apparatus

The stimuli were displayed on a computer-controlled Barco color monitor (Type HIREM). The chromaticities of generated lights were measured with the Minolta Chroma Meter CS100 and were checked for constancy at the beginning of each session. The set of 15 colors used in the experiments is shown in Figure 1. The stimuli were selected in such a way that, when presented to normal observers, they were separated by approximately equal subjective distances. In addition, subsets of the stimuli were chosen to lie along protanopic (G-GY-Y-O-R) and deuteranopic (GB-W-PR and BP-P) confusion lines, which also determine the long axes of discrimination ellipses of the corresponding types of anomalous trichromats (Regan, Reffin, & Mollon, 1994).

Pairs of colors were presented as 2° equiluminant (30 cd/m²) square patches in quasi-random order; the inner edge of each square was 1° from the fixation point. The background was dark. Over the course of months, the color display showed small variations in the functions relating the luminance outputs of the three guns to the signals submitted to the computer; therefore, fresh calibrations were made before each series of measurements for a given observer.

Procedure

In the RT experiment, on half the trials, the two stimulus patches were identical in color; each color could occur with equal frequency. On the remaining half of the trials, the stimulus patches were of different color. *Same* and *different* trials were randomly intermixed. The subject was required to press one of two keys—one for *same* and another for *different*—as rapidly as possible but was instructed to keep incorrect responses to a minimum. The colors disappeared when the key was pressed, and feedback (a buzzer) was given on incorrect responses. A new pair of colors appeared 2 sec after the response. The *same* trials served to oblige the subject to make a true discrimination and were not used in the subsequent analysis. Medians of *different* RTs, instead of means, were calculated, since RT distributions are known to be skewed (Boynton, Fargo, Olson, & Smallman, 1989). Erroneous "same" responses were excluded from the analysis. The resulting values for each subject contained the medians of all the RTs for a given pair of different colors.

In the rating experiment, each pair was presented for 0.5 sec. The task of an observer was to judge dissimilarity between colors on a 10-point scale (0–9, with 0 indicating *indistinguishable* and 9 indicating *maximal subjective dissimilarity*) and to name it to the experimenter, who entered the rating into the computer. After the input, the next pair of stimuli was presented automatically. The resulting values for each subject were the means of all the ratings for a given pair of colors.

Each color pair was presented 10 times in both the RT and the rating experiments.

Data Processing

To derive color spaces, the nonmetric MDS program KYST (Kruskal & Young, 1971; Multidimensional Scaling Package, 1989) was used, since we did not wish to assume any particular relationship, other than the minimal requirement that the distance between colors in a pair increases monotonically as RT decreases or as rated dissimilarity increases. For this study, three-, two-, and one-dimensional solutions were computed.

Figure 2. Scatter diagrams of dissimilarity ratings versus reaction times of *same-different* judgments of pairs of color stimuli for normal trichromats (NTs), protanomal siblings (PAs) H.A.N. and J.O.N., and deuteranope (D) E.H.R. The diagrams are accompanied by Spearman correlation coefficients (r_s), individual approximation exponential functions, and indices of variance (R^2).

RESULTS

Comparison of Reaction Time and Rating Measures of Color Difference¹

Absolute values of RTs for the NTs (330–700 msec) were typically lower than those for the color-abnormal observers: H.A.N. (420–1,800 msec), J.O.N. (400–

1,200 msec), and E.H.R. (470–1,400 msec). The number of incorrect *same* judgments given by the color-abnormal subjects (49–124) was much higher than that given by the normals (2–34). Erroneous judgments occurred primarily on color pairs that lie on the corresponding confusion lines, with some of these pairs being judged incorrectly by J.O.N. and E.H.R. on each of 10 presentations.

Table 1
Values of Stress for Solutions of Different Dimensionality
in Reaction Time and Rating Conditions for the
Three Normal Trichromats (NTs), the Two
Protanomal Siblings (PAs), and the Deuteranope (D)

Observer	Experimental Condition					
	Reaction Time			Rating		
	1-D	2-D	3-D	1-D	2-D	3-D
NT C.R.C.	.342	.128	.076	.285	.087	.053
NT E.L.E.	.362	.175	.114	.253	.073	.047
NT G.V.P.	.401	.205	.139	.334	.081	.062
PA H.A.N.	.413	.221	.137	.304	.088	.049
PA J.O.N.	.310	.158	.105	.214	.083	.037
D E.H.R.	.353	.183	.105	.085	.058	.038

Our finding is consistent with that of Taimela (1991), who found that subjects who made errors on the Ishihara test had significantly slower choice RTs and higher error rates in a task that required responding to colored lights by pressing corresponding colored keys. It should be noted that incorrect *same* responses in our study could not be used in the following MDS analysis, and this resulted in missing entries in the matrices for J.O.N. and E.H.R. (see Multidimensional Scaling Package, 1989).

Dissimilarity ratings show the expected inverse relationship to RTs. Abnormal rating matrices differ from those of the normals, in that there are more low and zero ratings ascribed to the pairs that were found to be highly confused in the RT data.

Values of ratings versus RTs are presented for each subject with scatter diagrams (Figure 2). The two sets of data are monotonically related, with point distributions fitted by the following exponential function:

$$y = B + (A - B)\exp(-kx),$$

where x is the dissimilarity rating between colors, y is the discriminative RT (in milliseconds), and A , B , and k are subject-specific constants. B indicates an approximation of the asymptotic minimum values of individual RTs, A is an approximation of the upper limit of the individual RT range, and k is the slope of the curve. On each diagram, approximating functions are presented, as well as R^2 , which evaluates the variance of the original distribution accounted for by the fitting function (in percentages). An inspection of individual scatter diagrams shows that even a very small inaccuracy in color discrimination, as diagnosed by traditional tests, reveals itself in a range of longer correct RTs and a higher variability of responses (and, thus, lower R^2); this tendency is exaggerated in color-deficient observers.

A nonparametric analysis (Spearman correlation coefficient, r_s) relating ratings and RTs for each subject was also performed. Individual r_s indices (presented on the diagrams) are comparable for the NTs and the color-deficient subjects. They do not, however, reveal the fact, which is obvious from visual inspection of the diagrams, that RTs discriminate small interstimulus differences well and are rather insensitive to larger differences. When color

differences are medium or large, they are evaluated differentially only by rating measures, since RT measures asymptote.

Color Spaces Reconstructed From Reaction Time and Rating Measures

Individual color spaces were reconstructed from each of the 12 matrices (2 per subject). Stress values were used as a formal measure of goodness-of-fit (see Table 1). Three dimensions were retained for further analysis. For the RT solutions, the values of stress (*fair-to-poor*, according to Kruskal, 1964) are higher than those for rating solutions (*good-to-fair*), suggesting that the RT data are rather "noisy." The higher stress in RT solutions is presumably due to there being only 10 presentations of each stimulus pair, for with 40 repetitions, in the studies of Cavonius and Mollon (1984; Müller et al., 1991), stress values were 0.08–0.09 and 0.04–0.06 for two- and three-dimensional solutions, respectively (personal communication).

Normal trichromats. Figures 3A–5A show color spaces reconstructed from RT measures for the NTs. Projected on the $D1$ – $D2$ plane, they closely resemble a conventional color circle, except that the distances in the blue–purple region are compressed, with white being displaced from the origin toward the blue–purple lights. An inspection of scatter diagrams relating RTs and reconstructed distances revealed that the longest RTs (and hence, the shortest distances) corresponded to discrimination of all the color pairs lying on tritanopic confusion lines—for example, I–BP, GB–G, PR–R, BP–W. Configurations projected on the $D1$ – $D2$ plane (with that for C.R.C. rotated by ca. -45°) can be reasonably interpreted in red–green and violet–yellow (tritanopic) dimensions.

In order to check whether the third dimension may reflect saturation (as was considered by Izmailov & Sokolov, 1991), we calculated Pearson correlation coefficients between the $D3$ coordinates of each color stimulus and the distance of the corresponding color from white in the CIE $u'v'$ chromaticity diagram (which is more uniform than the CIE x, y diagram; see Table 2). Significant correlation emerged for E.L.E.'s space, but not for the solutions for C.R.C. and G.V.P.

The rating color spaces (Figures 3B–5B), when plotted on the $D1$ – $D2$ plane, reveal a structure similar to the RT solutions. The former, however, are more regular, in the sense that they are not contracted in the blue–purple region and white occupies a position close to the origin in all three configurations. The axes rotated by ca. 45° can be also interpreted as red–green and violet–yellow dimensions. $D3$ is significantly correlated with distance from white in the CIE $u'v'$ chromaticity diagram only in C.R.C.'s solution.

Protanomals. In the RT spaces of the PAs, the $D1$ – $D2$ projection reveals a clustering of reds, greens, and yellow that lie on the protanopic confusion line in the CIE diagram. $D1$ opposes this cluster to blues and purples, the latter being discriminated much better, as is reflected by

Table 2
Pearson Correlation Coefficients Between CIE $u'v'$ Distance of Each Color Stimulus From White and Coordinates of Corresponding Colors on Conjectural Saturation Dimension

Observer	Dimension		Correlation Coefficient		$p <$	
	RT	Rating	RT	Rating	RT	Rating
NT C.R.C.	<i>D3</i>	<i>D3</i>	.12	-.70	.672	.004
NT E.L.E.	<i>D3</i>	<i>D3</i>	-.77	.12	.001	.671
NT G.V.P.	<i>D3</i>	<i>D3</i>	.30	-.19	.279	.490
NT average	<i>D3</i>	<i>D3</i>	-.64	-.83	.010	.000
PA H.A.N.	<i>D2</i>	<i>D3</i>	-.82	-.74	.000	.002
PA J.O.N.	<i>D2</i>	<i>D3</i>	-.69*	-.85	.009*	.000
D.E.H.R.	<i>D2</i>	<i>D3</i>	-.53	-.73	.041	.002

*Calculated for 13 distances, excluding GY and GB, whose coordinates due to empty cells were wrongly estimated by KYST. Bold symbols designate significant correlations. NT, normal trichromat; PA, protanomal; D, deuteranope.

the larger distances between them, with RTs being shorter for pairs on the tritanopic than for those on the protanopic direction. In H.A.N.'s configuration (Figure 6A), *D2* opposes indigo, on the one hand, to white and, on the other hand, to green-blue (which, on 7 out of 10 trials, was confused with white). In J.O.N.'s configuration, similar to his brother's, white and green-blue should logically fall at the same point of the space (in Figure 7A, shown by bold letters accompanied by arrows), for he responded *same* on all presentations. (The resulting RT cell was therefore empty, and the obtained separation between the two colors was estimated by KYST on the basis of the discriminability of all the other color pairs, yielding an implausible solution.) As is shown by significant correlations of *D2* coordinates with distances from white (Table 2), in the PAs' spaces, this dimension appears to reflect achromatic discrimination. Finally, *D3* clearly reveals the opposition of the red and the green lights and reflects the reduced red-green discrimination in these observers.

The rating *D1-D2* configurations of the PAs (Figures 6B and 7B) are similar to *D1-D2* configurations of the NTs, with *D2* reflecting red-green discrimination. They also show larger distances between red-yellow-green colors than in the corresponding RT solutions, in particular between green and green-yellow (which, under *same-different* condition, were confused by H.A.N. 7 times and by J.O.N. 10 times out of 10 presentations). *D3* lends itself to interpretation as saturation (see Table 2).

Deuteranope. In the RT solution projected on *D1-D2* (Figure 8A), the red-yellow-green colors lying on the deuteranopic confusion line in the CIE diagram also cluster and are opposed to blue-purple. The RTs to the pairs on the tritanopic direction were found to be the shortest, as compared with the RTs to the pairs on the orthogonal direction. The second dimension, *D2*, appears to be interpretable as saturation, being significantly correlated with a degree of stimulus whiteness (see Table 2), whereas *D3*, as would be expected in a dichromat, reveals no red-green discrimination.

The rating configuration (Figure 8B) generally resembles that in the RT-derived space: *D1* opposes the clus-

tered red-yellow-greens and the much better discriminated blue-purple. Again, *D2* reveals discrimination of saturation, and *D3* is redundant.

DISCUSSION

The main purpose of the present study was to determine whether the similarity structure derived from judgments of color differences is dependent on which of two experimental conditions that differ in their nature (discrimination vs. dissimilarity) and mode (speeded vs. unspeeded) of response is used, even though the subjects, equipment, and stimuli are exactly the same in all respects. In addition, an observer factor (normal vs. deficient color processing) was examined. The goal was to compare measures of color differences obtained under the two conditions and the spatial solutions derived from them (configurations, number of dimensions, and order of their significance).

Comparison of Reaction Time and Dissimilarity Measures

Discriminative RTs in a *same-different* task were found to decline exponentially as dissimilarity ratings increased. This indicates that, for a color set comprising both small and large color differences, RTs are sensitive only in near-threshold region, where they are related to detectability of color differences that are slightly above discrimination threshold (see Luce, 1986). As dissimilarity increases, RTs decline toward an asymptotic minimum, and the rank order among them is lost in the noise. This has a detrimental outcome when submitting matrices with such properties to MDS analysis: The MDS algorithm tends to equalize large distances while still providing good fits at shorter distances. This leads to a displacement of points in the RT-derived space and curling of the representation of the points into new dimensions, an artifact resulting from disordered RT ranking (for comments, see Bimler, 1999).

Comparison of Spatial Solutions Derived From Reaction Times and Ratings

Stress values. A formal comparison of spatial representations based on different data collection procedures can be provided by *stress*, an index of goodness of fit for each solution. Higher stress values for RT solutions, as compared with those for rating ones, indicate that the *same-different* paradigm produces more *noise*, although when the number of trials in the latter is increased two to three times, the stress values for both types of solutions become comparable (cf. Young, 1970).

Apart from the noisier data, applying MDS to RT measures obtained from color-deficient observers has the further disadvantage that incorrect *same* responses have to be discarded and, therefore, a good deal of information on color discrimination is lost, especially on pairs that are invariably confused. A suggestion for solving this methodological problem is to process matrices consisting of the proportion of correct responses, in addition to RT

Figure 3. Color spaces for normal trichromat C.R.C., reconstructed from reaction times (A) and ratings (B): projections of the configurations onto the $D1$ - $D2$ plane (left panels) and the $D1$ - $D3$ plane (right panels).

measures (see Young, 1970). In the present study, however, this was found to be impractical, for a significant number of color pairs were suprathreshold and, thus, generally produced correct responses, leading to MDS solutions that appeared to be degenerate.

Dimensionality and configuration. Against our expectations, RT- and rating-derived color spaces showed no consistent differences between the dimensional outcomes of the two procedures, although they indicated individual variations in dimensional solutions.

For NTs under each condition, two chromatic dimensions were found (red-green and violet-yellow), indicating that dissimilarity judgment is governed by the same level of color processing as is discrimination of colors. In addition, a third dimension emerged in the RT space

of E.L.E. and the rating space of C.R.C., although its contribution was rather small. It is worth noting that, in the RT- and rating-derived solutions, averaged for the three NTs (not shown here), the third dimension also emerged (see Table 2). As Ashby, Maddox, and Lee (1994) argued, this might be due to the increased symmetry that generally results from averaging data and reveals trends that are less clear in individual solutions, presumably because the latter contain more noise. The averaging of data (of 15 subjects) may be the reason that the third dimension emerged in the study of Sokolov and Izmailov (1983).

We are aware that representation of equiluminant lights in three dimensions is unconventional. The emergence of the additional dimension becomes more clear if one takes

Figure 4. Color spaces for normal trichromat E.L.E., reconstructed from reaction times (A) and ratings (B): projections of the configurations onto the $D1$ – $D2$ plane (left panels) and the $D1$ – $D3$ plane (right panels).

into consideration the fact that the three-dimensional configurations for the NTs (as well as those for the PAs, described below) were found to lie on a spherical surface (see Bimler, 1999). This corroborates the observation that the color plane is curved, made by Judd (1966) for equi-light colors and by Sokolov and Izmailov (1983) for equi-luminant lights. When mapped in an Euclidean space, a curved configuration produces an additional dimension (Blough, 1988). As Bimler notes, the third dimension in color space can be tailored to the explanation that the evaluation of color differences includes contributions from a separate quality of color. An argument for the nonartificial nature of $D3$ is its high correlation with the degree of stimulus whiteness, which is consistent with the labeling of this dimension as saturation by Izmailov and Sokolov (1991), especially in view of the fact that a saturation di-

mension emerges when stimuli in a color set vary substantially in purity (see Sergent & Takane, 1987). Despite this straightforward interpretation, it should be noted that saturation could also have been confounded with brightness (Helmholtz–Kohlrausch effect), for when colors are equated in luminance, the more saturated of the two colors appears to be brighter (Uchikawa, Uchikawa, & Kaiser, 1984).

We found consistent differences between the configurations derived for the NTs from the two measures. The rating configurations indicate a more regular location of points along the “color circle,” whereas the configurations in RT-derived color spaces reveal contraction of distances along the tritanopic direction. The latter finding conforms to that of Mollon and Cavonius (1986), who showed that stimuli equally spaced in the CIE $u'v'$ chro-

Figure 5. Color spaces for normal trichromat G.V.P., reconstructed from reaction times (A) and ratings (B): projections of the configurations onto the $D1$ - $D2$ plane (left panels) and the $D1$ - $D3$ plane (right panels).

maticity diagram, when presented in a RT-derived color space, yield a configuration that is contracted in the violet–yellow direction, corresponding to one of the cardinal directions of Krauskopf et al. (1982).

The solutions for the red–green-deficient observers are dominated, as would be expected, by the violet–yellow dimension. In accordance with the previous finding of Paramei et al. (1991), there is evidence for a saturation dimension, implying that saturation clues appear to be more salient for these observers, compensating for deficiencies in red–green processing.

For the PAs, both RT- and rating-derived solutions are three-dimensional. The discrepancy between the PAs' solutions obtained under the two conditions concerns the order of significance of the red–green dimension—

namely, $D3$ in the RT- versus $D2$ in the rating-spaces. The saturation dimension ($D2$ in the RT- and $D3$ in the rating-spaces) may be more confounded with brightness, for protans without italics are expected to demonstrate greater variation of the brightness/luminance ratio than do NTs, because of their lower sensitivity to long-wavelength lights (Hurvich, 1981). The interchange of the order of significance indicates that, when measured by RTs, red–green discrimination appears to be less effective than discrimination of saturation.

We would like to emphasize that the PAs' spaces in the present study reveal a red–green dimension that is consistent with the previous (dis)similarity findings (Helm, 1964; Paramei, 1982), but at variance with the absence of a red–green dimension found in RT solutions (Cavonius

Figure 6. Color spaces for protanomal H.A.N., reconstructed from reaction times (A) and ratings (B): projections of the configurations onto the $D1$ - $D2$ plane (left panels) and the $D1$ - $D3$ plane (right panels).

et al., 1990). As was mentioned in the introduction, the latter authors supposed that the second dimension might reflect saturation of equiluminant stimuli. This is in agreement with our finding that the second dimension in RT spaces reflects saturation. It is conceivable that retention of only two dimensions in the study of Cavanaugh et al. might lead to a "loss" of the red-green dimension.

Finally, the RT-derived configurations for the PAs, when compared with the rating ones, indicate clustering of lights lying on the protanopic confusion line and, in addition, contraction of distances in the tritanopic direction (as in the NTs).

For the deuteranope E.H.R., two-dimensional solutions obtained under both experimental procedures show no

principal differences, apart from an apparent, presumably degenerate, horseshoe configuration in the RT space.

What May Account for the Differences

Our first consideration, in accounting for the discrepancies between the solutions derived from the RTs and those from the dissimilarity judgments, was that they are due to the response mode: fast for RT trials versus moderate for ratings. The data for the NTs suggest that the longer stimulus duration in the rating experiment (500 msec) provides more information than does the limited viewing time in the RT experiment: If we subtract the motor component of the RT, the sensory component would be somewhere in the range of 150–400 msec. This

consideration is consistent with the finding that hue discrimination thresholds rise dramatically at short—for example, below ca. 250 msec—exposure durations (Rütiger, Kurtenbach, & Zrenner, 1995) or even below 1 sec (Hita, Romero, Jiménez del Barco, & Martínez, 1982).

A more likely explanation concerns the effect of attention on processing time (Palmer, Ames, & Lindsey, 1993), which shows itself in RTs as measures of color differences. If a pair of colors must be examined along all dimensions before initiating the response *different* (Egeth, 1966) and a certain chromatic mechanism has a relatively long integration time, color discrimination along it and, therefore, the overall discrimination would take longer. Indeed, our data

show that, in NTs, discrimination of color pairs that lie on tritanopic confusion lines yielded longer RTs. Since tritan discriminations require inputs from the short-wavelength cones, this agrees with the finding that latencies of S-cones are greater than those of L- or M-cones (Mollon & Krauskopf, 1973). In addition, there may be later, related, neural factors: It has been claimed that the temporal resolution of the S/(L+M) pathway is about 30% slower than that of the M/L pathway (Courtney & Buchsbaum, 1991). Both of these findings are consistent with the compression along tritanopic confusion lines in our results.

In contrast, in the color-deficient observers, RTs to color pairs lying on protanopic or deuteranopic confusion lines

Figure 8. Color spaces for deuteranope E.H.R., reconstructed from reaction times (A) and ratings (B): projections of the configurations onto the $D1$ - $D2$ plane (left panels) and the $D1$ - $D3$ plane (right panels).

are longer than RTs to those in the tritanopic direction, which results in general lengthening of color processing.

The present findings show that, when using different measures for the modeling of color spaces, one should take account of the specific tasks and measures from which the spaces are derived.

REFERENCES

- ASHBY, F. G., MADDOX, W. T., & LEE, W. W. (1994). On the dangers of averaging across subjects when using multidimensional scaling or the similarity-choice model. *Psychological Science*, *5*, 144-151.
- BIMLER, D. (1999). Research note. A multidimensional scaling comparison of color metrics for response times and rated dissimilarities. *Perception & Psychophysics*.
- BLOUGH, D. S. (1988). Quantitative relations between visual search speed and target-distractor similarity. *Perception & Psychophysics*, *43*, 57-71.
- BORG, I., & LINGOES, J. (1987). *Multidimensional similarity structure analysis*. New York: Springer-Verlag.
- BOYNTON, R. M., FARGO, L., OLSON, C. X., & SMALLMAN, H. S. (1989). Category effects in color memory. *Color Research & Application*, *14*, 229-234.
- CAVONIUS, C. R., & MOLLON, J. D. (1984). Reaction time as a measure of the discriminability of large color differences. In C. P. Gibson (Ed.), *Proceedings of a Workshop on Colour Coded vs Monochrome Electronic Displays* (NATO Workshop, Farnborough, England, 28.2.-1.3.1984; pp. 17.1-17.10). London: HMSO.
- CAVONIUS, C. R., MÜLLER, M., & MOLLON, J. D. (1990). Difficulties faced by color-anomalous observers in interpreting color displays. *SPIE Proceedings*, *1250*, 190-195.
- CHANG, J. J., & CARROLL, J. D. (1980). Three are not enough: An INDSCAL analysis suggesting that color space has seven (± 1) dimensions. *Color Research & Application*, *5*, 193-206.

- COURTNEY, S. M., & BUCHSBAUM, G. (1991). Temporal differences between color pathways within retina as a possible origin of subjective colors. *Vision Research*, **31**, 1541-1548.
- EGETH, H. F. (1966). Parallel versus serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics*, **1**, 245-252.
- EKMAN, G. (1954). Dimensions of color vision. *Journal of Psychology*, **38**, 467-474.
- HELM, C. E. (1964). A multidimensional ratio scaling analysis of perceived color relations. *Journal of the Optical Society of America*, **54**, 256-262.
- HELM, C. E., & TUCKER, R. (1962). Individual differences in the structure of color perception. *American Journal of Psychology*, **75**, 437-444.
- HITA, E., ROMERO, J., JIMÉNEZ DEL BARCO, L., & MARTÍNEZ, R. (1982). Temporal aspects of color discrimination. *Journal of the Optical Society of America*, **72**, 578-582.
- HURVICH, L. M. (1981). *Color vision*. Sunderland, MA: Sinauer.
- IZMAILOV, C. A., & SOKOLOV, E. N. (1991). Spherical model of color and brightness discrimination. *Psychological Science*, **2**, 249-259.
- JORDAN, G., & MOLLON, J. D. (1993). A study of women heterozygous for colour deficiencies. *Vision Research*, **33**, 1495-1508.
- JUDD, D. B. (1966). Progress report for OSA Committee on Uniform Color Scales. In *Proceedings of the International Colour Meeting* (pp. 399-407). Göttingen: Musterschmidt.
- KRAUSKOPF, J., WILLIAMS, D. R., & HEELEY, D. W. (1982). Cardinal directions of color space. *Vision Research*, **22**, 1123-1131.
- KRUSKAL, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, **29**, 1-27.
- KRUSKAL, J. B., & YOUNG, F. W. (1971). *Kruskal-Young-Shepard and Torgerson (KYST) multidimensional scaling program: A Fortran IV program*. Murray Hill, NJ: AT&T Bell Laboratories.
- LUCE, R. D. (1986). *Response times. Their role in inferring elementary mental organization*. Oxford: Oxford University Press.
- MOLLON, J. D., & CAVONIUS, C. R. (1986). The discriminability of colours on c.r.t. displays. *Journal of the Institution of Electronic & Radio Engineers*, **56**, 107-110.
- MOLLON, J. D., & KRAUSKOPF, J. (1973). Reaction time as a measure of the temporal response properties of individual colour mechanisms. *Vision Research*, **13**, 27-40.
- MÜLLER, M., CAVONIUS, C. R., & MOLLON, J. D. (1991). Constructing the color space of the deuteranomalous observer. In B. Drum, J. D. Moreland, & A. Serra (Eds.), *Colour vision deficiencies X* (pp. 377-387). Dordrecht: Kluwer.
- Multidimensional scaling package* (1989). Murray Hill, NJ: AT&T Bell Laboratories.
- PALMER, J., AMES, C. T., & LINDSEY, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology. Human Perception & Performance*, **19**, 108-130.
- PARAMEI, G. V. (1982). Características cuantitativas de las alteraciones de la visión de los colores [Quantitative characteristics of color vision alterations]. *Boletín de Psicología*, **5**, 67-91.
- PARAMEI, G. V., IZMAILOV, C. A., & SOKOLOV, E. N. (1991). Multidimensional scaling of large chromatic differences by normal and color-deficient subjects. *Psychological Science*, **2**, 244-248.
- REGAN, B. C., REFFIN, J. P., & MOLLON, J. D. (1994). Luminance noise and the rapid determination of discrimination ellipses in colour deficiency. *Vision Research*, **34**, 1279-1299.
- RÜTTIGER, L., KURTENBACH, A., & ZRENNER, E. (1995). Exposure duration, intensity-time relationship, and stimulus onset asynchrony affect wavelength-discrimination ability [Abstract]. *Perception*, **24**(Suppl.), 55.
- SERGEANT, J., & TAKANE, Y. (1987). Structures in two-choice reaction-time data. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 300-315.
- SHEPARD, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function: II. *Psychometrika*, **27**, 219-246.
- SHEPARD, R. N. (1980). Multidimensional scaling, tree fitting and clustering. *Science*, **210**, 390-398.
- SHEPARD, R. N. (1986). Discrimination and generalization in identification and classification: Comment on Nosofsky. *Journal of Experimental Psychology: General*, **115**, 58-61.
- SOKOLOV, E. N., & IZMAILOV, C. A. (1983). The conceptual reflex arc: A model of neural processing as developed for colour vision. In H.-G. Geissler (Ed.), *Modern issues in perception* (pp. 192-216). Amsterdam: North-Holland.
- TAIMELA, S. (1991). Factors affecting reaction-time testing and the interpretation of results. *Perceptual & Motor Skills*, **73**, 1195-1202.
- TANSLEY, B. W., & BOYNTON, R. B. (1976). A line, not a space, represents visual distinctness of borders formed by different colors. *Science*, **191**, 954-957.
- UCHIKAWA, K., UCHIKAWA, H., & KAISER, P. (1984). Luminance and saturation of equally bright colors. *Color Research & Application*, **9**, 5-14.
- YOUNG, F. W. (1970). Nonmetric scaling of line lengths using latencies, similarity, and same-different judgments. *Perception & Psychophysics*, **8**, 363-369.
- YOUNG, F. W., & HAMER, R. M. (1987). *Multidimensional scaling: History, theory, and applications*. Hillsdale, NJ: Erlbaum.

NOTE

1. Individual matrices of the discriminative RTs and dissimilarity measures are available from the authors.

(Manuscript received August 19, 1996;
revision accepted for publication July 27, 1998.)