4

Three Applications of MDS

Three applications of MDS are discussed in some depth. Emphasis is given to the questions of how to choose a particular MDS solution and how to interpret it. First, data on the perceived similarity of colors are studied. The predicted MDS configuration is a color circle, which is indeed found to be the best representation for the data. Second, confusion data on Morse codes are investigated. The MDS space shows two regional patterns, which reflect two physical properties of the signals. Third, global similarity judgments on different facial expressions are studied. A dimensional system can be found that relates to three empirical scales for the faces.

4.1 The Circular Structure of Color Similarities

We now look at some applications of MDS in somewhat more depth and not just in an illustrative way. We start with a classic case where the MDS solution is particularly revealing.

Some Data on the Perceived Similarity of Colors

A person asked to somehow orderly arrange chips of different colors will almost certainly come up with an order from orange over yellow, green, blue, to blue-violet, corresponding to the order of the electromagnetic wavelengths of these colors. For the color red-violet, the respondent would probably not be sure whether it should lie on the red end or the violet end of the scale (or on both). This problem is solved by arranging the colors in a horseshoe or circle. It may be supposed that most persons confronted with this ordering task would sooner or later arrive at such a solution.

Both perceptual and intellectual components seem to be involved in solving this task, but to what relative extent? It could be argued that the color circle is already implied by the way we perceive the similarity of colors. Let us look at some data by Ekman (1954). Ekman used 14 colors differing only in their wavelengths, but not in their brightness or saturation. Each of all possible 91 pairs of different colors was projected onto a screen, and 31 subjects were asked to rate the "qualitative similarity" of each such pair on a scale from 0 (no similarity) to 4 (identical). The ratings for each pair were averaged over all subjects. Finally, the resulting scores were divided by 4, that is, scaled down to the interval from 0 to 1. This led to the similarity matrix in Table 4.1 (lower half). Note that only one-half of the matrix was collected empirically, and so it suffices to show this half: the complete matrix, if needed, can be constructed by setting $p_{ii} = 1.00$ and $p_{ij} = p_{ji}$, for all i, j. (Most MDS programs need only a half-matrix as input.)

The proximities in Table 4.1 could be interpreted as correlations, so that a principal component analysis (PCA; see also Chapter 24) is possible. A PCA yields five different factors. These factors correspond to five different groups of points on the electromagnetic spectrum. The factors comprise the colors 434–445, 465–490, 504–555, 584–600, and 610–674, which roughly correspond to the subjective color qualities blueish-purple, blue, green, yellow, and red, respectively. Chopping up the colors into qualitative categories, however, does not throw much light on the question we are asking.

An inspection of the coefficients in Table 4.1 shows that the data do not support the notion of discrete color categories. Rather, they possess a simple pattern of interrelatedness, a peculiar gradient of similarities, with larger coefficients towards the main diagonal and the lower left-hand corner, respectively. So, using MDS, which establishes a direct relationship between dissimilarity measures and geometric distance (unlike PCA), we would possibly get a simple geometric expression for this data gradient.

MDS Representations of the Color Similarities

For the MDS analysis, we use ordinal MDS, the usual choice for a first approximation. Thus, a configuration of 14 points is sought such that the rank-order of the distances between these points corresponds (inversely) to the rank-order of the data.

Any MDS program requires the user to specify the dimensionality (m) of the desired representation. What value m should be chosen in the given case? Surely, setting $m \geq 13$ would be an uninteresting choice, because dissimilarities among n objects can always be perfectly represented in a space with dimensionality $m \geq n - 1$. For example, in a plane with points

TABLE 4.1. Similarities of colors with wavelengths from 434 to 674 nm (lower half) of Ekman (1954); residuals of 1D MDS representation (upper half).

nm	434	445	465	472	490	504	537	555	584	600	610	628	651	674
434	–	.14	.17	.38	.22	-.73	-1.07	-1.21	-.62	-.06	.42	.38	.28	.26
445	.86	–	.25	.11	-.05	-.75	-1.09	-.68	-.35	-.04	.44	.65	.55	.53
465	.42	.50	–	.08	-.32	-.57	-.47	-.06	.00	-.32	.17	.12	.91	.82
472	.42	.44	.81	–	.12	-.36	-.26	.15	.00	-.11	.00	.33	.23	1.03
490	.18	.22	.47	.54	–	-.07	.08	.48	.40	.00	.22	.17	.07	.00
504	.06	.09	.17	.25	.61	–	.31	.28	.45	.68	.01	.00	.00	-.15
537	.07	.07	.10	.10	.31	.62	–	.13	.35	.09	.31	.00	.00	-.75
555	.04	.07	.08	.09	.26	.45	.73	–	-.05	.17	-.09	-.22	-.32	-.34
584	.02	.02	.02	.02	.07	.14	.22	.33	–	-.05	-.01	-.06	-.16	-.18
600	.07	.04	.01	.01	.02	.08	.14	.19	.58	–	.21	.07	-.39	-.40
610	.09	.07	.02	.00	.02	.02	.05	.04	.37	.74	–	-.08	-.13	-.11
628	.12	.11	.01	.01	.01	.02	.02	.03	.27	.50	.76	–	-.03	-.16
651	.13	.13	.05	.02	.02	.02	.02	.02	.20	.41	.62	.85	–	-.11
674	.16	.14	.03	.04	.00	.01	.00	.02	.23	.28	.55	.68	.76	–

A, B, and C, it is possible, by moving the points around, to eventually arrive at a configuration whose distances perfectly represent any given proximities $p(A, B)$, $p(A, C)$, $p(B, C)$, no matter what values they have. Analogously, for four points, a perfect representation always exists in three dimensions, and so on.

The minimal dimensionality for a perfect MDS representation is only of formal interest. In practice, we always try to represent the data in an MDS space of considerably lower dimensionality. The rationale for choosing a low dimensionality is the expectation that this will cancel out over- and under-estimation errors in the proximities, thus smoothing the representation (see Chapter 3). Moreover, a low-dimensional and preferably two-dimensional solution is often precise enough for a first interpretation. For the data in Table 4.1, we first try solutions in 1D, 2D, and 3D space (using the MDS module of SYSTAT 5.0).

With three solutions, we have to decide which one we should consider most appropriate. We first look at the 2D solution in Figure 4.1. It shows a circular arrangement of the points representing the colors. Moreover, the points are perfectly ordered along the drawn-in line in terms of their wavelengths. This circular structure corresponds to the color circle.

How well does this configuration represent the data? The best answer to this question is provided by looking at the Shepard diagram of the 2D MDS solution (Figure 4.2). The plot shows a tight correspondence of proximities and distances. The points lie very close to the monotone regression line. The regression line is almost straight, and so the dissimilarities of Table 4.1 are almost linearly and almost perfectly related to the distances in Figure 4.1. In contrast, in the Shepard diagram for the 1D solution (Figure

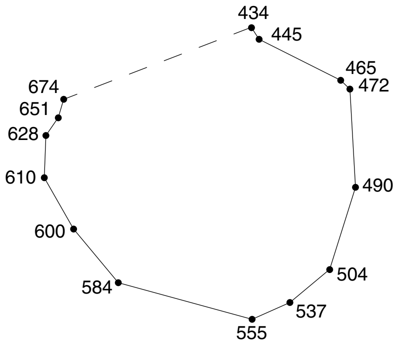

FIGURE 4.1. Ordinal MDS representation for color proximities in Table 4.1.

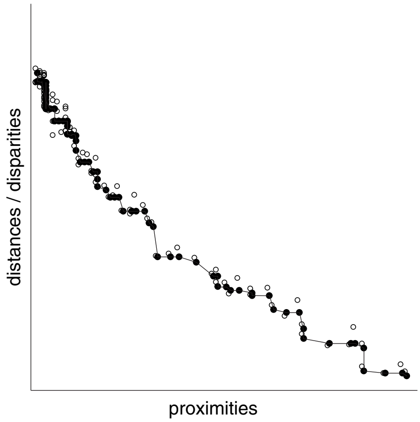

FIGURE 4.2. Shepard diagram for Fig. 4.1.

4.4), the deviations of the points from the shown best-possible monotonic decreasing line are excessive.

Measured in terms of Stress, the badness-of-fit of the 1D, 2D, and 3D solutions is 0.272, 0.023, and 0.018, respectively. These values are a rare example for a definite elbow in the scree test. The 1D solution has high Stress, and adding one additional dimension leads to a major Stress reduction. Adding yet another dimension has very little further effect and, indeed, cannot have much of an effect because the 0.023 for the 2D solution is so close to zero already.

Thus, the 2D solution appears to be a reasonably precise representation of the data. Adding a third dimension is not sensible, because of several reasons: (a) the point configuration in the X–Y-plane of the 3D solution (Figure 4.3) corresponds closely to the 2D configuration (Figure 4.1); (b) the decrement in Stress by allowing for a third dimension is negligible, satisfying the elbow criterion; (c) the scattering of the points in 3D space along the third dimension appears to be uninterpretable in substantive terms; and (d) no a priori theory exists for a 3D solution. Analogous arguments hold for comparing the 1D and 2D solutions. Hence, we have formal and substantive reasons to consider the 2D representation in Figure 4.1 as the best MDS representation of the given data.

A Closer Look at Model Fit

The Shepard diagram in Figure 4.4 shows that the 1D solution is a relatively poor representation of the data. Why there cannot exist a really good 1D solution can be seen from Figure 4.1. If we had to locate a straight line in this plane so that the distances between the projections of the points onto this line mirror most closely the order of the data, then this line would be

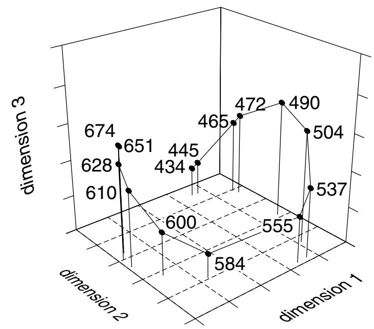

FIGURE 4.3. 3D MDS space of color data.

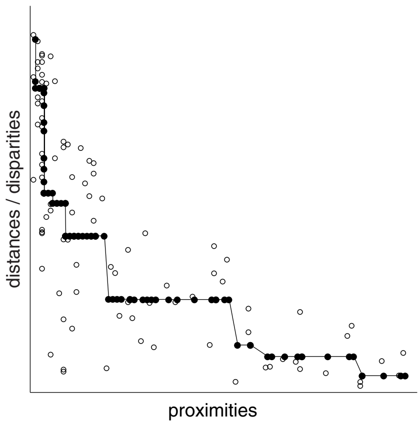

FIGURE 4.4. Shepard diagram for 1D MDS representation of color data.

oriented roughly horizontally. Such a line is the best 1D approximation to the given 2D distance structure, because most of the straight lines connecting any two points in Figure 4.1 run more or less in this direction. For point 610, for example, we see in Figure 4.1 that the projections of the rays from this point to all other points onto a horizontal line are ordered (in length) almost as the rays themselves. However, there would also be misrepresentations on this line. For example, the points 434 and 555, if projected onto a horizontal line, would be very close to each other, whereas the similarity between 434 and 555 is among the lowest ones observed. Hence, this datum is not represented well by the points' distance on this 1D subspace.

We should expect, then, that a 1D MDS solution for the color data represents the proximities of such colors as 610 and 472 with respect to all other colors quite well, but that it runs into problems with pairs such as 434 and 555. One could assess such effects quantitatively by computing, for each color C in turn, the correlation between the similarities of C to all other colors and the distances of point C to all other points. To be consistent with the ordinal approach, an ordinal correlation (e.g., Spearman's ρ) would be appropriate. Each such coefficient is a *conditional* fit measure, because it hinges on one fixed point or variable (C, here).

Using Spearman correlations, one finds that they are, for each point C, close to -1.00 for the 2D and 3D solutions. For the 1D case, in contrast, there is much more variance. The coefficients are particularly low for points 434 ($r = -.075$) and 445 ($r = -.360$) at the upper end of the horseshoe in Figure 4.1. Low conditional fit measures imply that the overall precision of the 1D MDS representation (as measured by Stress, e.g.) cannot be very good, because conditional agreements between distances and data are a necessary condition for a globally good solution.

Spearman's correlation is, however, not very robust, because it is based on the ranks of the data, and major changes of the rank-order sometimes result from minor changes of the data. A correlation coefficient that assesses the degree of monotone correspondence directly on the data is μ_2 (Guttman, 1968; see also Chapter 14). For the given data, however, one arrives at the same conclusion using μ_2: points 434 and 445 are the major sources of Stress.

An even more fine-grained analysis of the sources of Stress is possible by studying the residuals, e_{ij}, for all i, j. Table 4.1 (upper half) shows these residuals for the 1D MDS representation of the color data. One notes, for example, that the similarity measures for the pairs $(434, 555)$ and $(434, 537)$ are relatively poorly represented by their corresponding distances, as expected. For the pair $(610, 472)$, in contrast, the residual is zero.

Most MDS computer programs provide these residuals upon request. Some also compute some kind of average residual value — such as the root mean squared residual — for each point in turn. Such coefficients are conditional fit measures closely related to the Stress formula (Borg, 1978b).

4.2 The Regionality of Morse Codes Confusions

The next example we consider is also from perception on stimuli that are physically well structured. This most complex data matrix requires special considerations and some simplifications. The MDS configuration, then, clearly reflects the structural properties of the stimuli.

Morse Code Confusions and Their Representability by Distances

Consider now the data matrix in Table 4.2 (Rothkopf, 1957). The scores are confusion rates on 36 Morse code signals (26 for the alphabet; 10 for the numbers $0, \ldots, 9$). Each Morse code signal is a sequence of up to five "beeps." The beeps can be short (0.05 sec) or long (0.15 sec), and, when there are two or more beeps in a signal, they are separated by periods of silence (0.05 sec). For example, the signal for A is "short-silence-long," with a total temporal length of 0.25 seconds. We code such a signal as 12 (1 = short and 2 = long, or "di-da").

Rothkopf (1957) asked 598 subjects to judge whether two signals, presented acoustically one after another, were the same. The values given in Table 4.2 are the percentages with which the answer "Same!" was given in each combination of row stimulus i and column stimulus j, where i was the first and j the second signal presented. Each stimulus pair was presented in two orders, for example, B following A (confusion rate is 4%) and also A following B (5%). Moreover, the rate of confusion of each signal with itself was assessed. For example, the relative frequency of confusing A with itself is 92%, and for B, 84%.

If we attempt an MDS representation, we notice several problems. First, we observe that the nonnegativity axiom does not hold, because the values in the main diagonal of the data matrix are not all the same. Because the distance from any point to itself is always 0, we will therefore necessarily incur a misrepresentation of the empirical data in the MDS space. On the other hand, the second part of the nonnegativity axiom poses no problem, because all data values in the main diagonal are greater than any off-diagonal value, and this can be properly expressed by distances in an MDS space.

Then, we see that the symmetry condition [axiom (2.2)] also does not hold for the data. For example, the signal I is more frequently confused with a subsequent A (64%) than A is with a subsequent I (46%). But if we represent I and A by one point each, then we will necessarily have the relation $d_{IA} = d_{AI}$, so that the asymmetry of the observed relation is lost, that is, not represented.

Finally, the triangle inequality can be checked only if the data are on a ratio scale. For all weaker MDS models, it is always possible to find a constant k so that every $p_{ij}+k$ satisfies the triangle inequality. The minimal constant k is found by first identifying the triangle inequality violated most and then computing the value that, when added to each proximity in this inequality, turns the inequality into an equality. Thus, unless we consider the Rothkopf data as ratio-scaled distances (apart from error), axiom (2.3) is immaterial.

Distance axioms (2.1) and (2.2), on the other hand, remain violated even if one allows for ordinal transformations of the data. Yet, we should take into account that none of Rothkopf's subjects knew Morse codes. It is a very demanding task for an untrained subject to distinguish consistently between different signals, and we might, therefore, argue that the violations of these axioms are unsystematic and due to error. (We test this in Chapter 24.) Under this assumption, we can think of the data for each (i, j) and (j, i) pair as replicated observations of a basically symmetric relation, and then obtain a better estimate of the true relation by averaging the two observed values. In other words, from Table 4.2 we form a new proximity matrix, where, say, $p_{AB} = p_{BA} = (.05 + .04)/2 = .045$. The main diagonal could then be filled with the value 1.00, say, although this is immaterial, because MDS programs ignore these values anyway.

TABLE 4.2. Confusion percentages between Morse code signals (Rothkopf, 1957).

Morse Code		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6	7	8	9	0	
·—	A	92	4	6	13	3	14	10	13	46	5	22	3	25	34	6	6	9	35	23	6	37	13	17	12	7	3	2	7	5	5	8	6	5	6	2	3	A
—···	B	5	84	37	31	5	28	17	21	5	19	34	40	6	10	12	22	25	16	18	2	18	34	8	84	30	42	12	17	14	40	32	74	43	17	4	4	B
—·—·	C	4	38	87	17	4	29	13	7	11	19	24	35	14	3	9	51	34	24	14	6	6	11	14	32	82	38	13	15	31	14	10	30	28	24	18	12	C
—··	D	8	62	17	88	7	23	40	36	9	13	81	56	8	7	9	27	9	45	29	6	17	20	27	40	15	33	3	9	6	11	9	19	8	10	5	6	D
·	E	6	13	14	6	97	2	4	4	17	1	5	6	4	4	5	1	5	10	7	67	3	3	2	5	6	5	4	3	5	3	5	2	4	2	3	3	E
··—·	F	4	51	33	19	2	90	10	29	5	33	16	50	7	6	10	42	12	35	14	2	21	27	25	19	27	13	8	16	47	25	26	24	21	5	5	5	F
——·	G	9	18	27	38	1	14	90	6	5	22	33	16	14	13	62	52	23	21	5	3	15	14	32	21	23	39	15	14	5	10	4	10	17	23	20	11	G
····	H	3	45	23	25	9	32	8	87	10	10	9	29	5	8	8	14	8	17	37	4	36	59	9	33	14	11	3	9	15	43	70	35	17	4	3	3	H
··	I	64	7	7	13	10	8	6	12	93	3	5	16	13	30	7	3	5	19	35	16	10	5	8	2	5	7	2	5	8	9	6	8	5	2	4	5	I
·———	J	7	9	38	9	2	24	18	5	4	85	22	31	8	3	21	63	47	11	2	7	9	9	9	22	32	28	67	66	33	15	7	11	28	29	26	23	J
—·—	K	5	24	38	73	1	17	25	11	5	27	91	33	10	12	31	14	31	22	2	2	23	17	33	63	16	18	5	9	17	8	8	18	14	13	5	6	K
·—··	L	2	69	43	45	10	24	12	26	9	30	27	86	6	2	9	37	36	28	12	5	16	19	20	31	25	59	12	13	17	15	26	29	36	16	7	3	L
——	M	24	12	5	14	7	17	29	8	8	11	23	8	96	62	11	10	15	20	7	9	13	4	21	9	18	8	5	7	6	6	5	7	11	7	10	4	M
—·	N	31	4	13	30	8	12	10	16	13	3	16	8	59	93	5	9	5	28	12	10	16	4	12	4	16	11	5	2	3	4	4	6	2	2	10	2	N
———	O	7	7	20	6	5	9	76	7	2	39	26	10	4	8	86	37	35	10	3	4	11	14	25	35	27	27	19	17	7	7	6	18	14	11	20	12	O
·——·	P	5	22	33	12	5	36	22	12	3	78	14	46	5	6	21	83	43	23	9	4	12	19	19	19	41	30	34	44	24	11	15	17	24	23	25	13	P
——·—	Q	8	20	38	11	4	15	10	5	2	27	23	26	7	6	22	51	91	11	2	3	6	14	12	37	50	63	34	32	17	12	9	27	40	58	37	24	Q
·—·	R	13	14	16	23	5	34	26	15	7	12	21	33	14	12	12	29	8	87	16	2	23	23	62	14	12	13	7	10	13	4	7	12	7	9	1	2	R
···	S	17	24	5	30	11	26	5	59	16	3	13	10	5	17	6	6	3	18	96	9	56	24	12	10	6	7	8	2	2	15	28	9	5	5	5	2	S
—	T	13	10	1	5	46	3	6	6	14	6	14	7	6	5	6	11	4	4	7	96	8	5	4	2	2	6	5	5	3	3	3	8	7	6	14	6	T
··—	U	14	29	12	32	4	32	11	34	21	7	44	32	11	13	6	20	12	40	51	6	93	57	34	17	9	11	6	6	16	34	10	9	9	7	4	3	U
···—	V	5	17	24	16	9	29	6	39	5	11	26	43	4	1	9	17	10	17	11	6	32	92	17	57	35	10	10	14	28	79	44	36	25	10	1	5	V
·——	W	9	21	30	22	9	36	25	15	4	25	29	18	15	6	26	20	25	61	12	4	19	20	86	22	25	22	10	22	19	16	5	9	11	6	3	7	W
—··—	X	7	64	45	19	3	28	11	6	1	35	50	42	10	8	24	32	61	10	12	3	12	17	21	91	48	26	12	20	24	27	16	57	29	16	17	6	X
—·——	Y	9	23	62	15	4	26	22	9	1	30	12	14	5	6	14	30	52	5	7	4	6	13	21	44	86	23	26	44	40	15	11	26	22	33	23	16	Y
——··	Z	3	46	45	18	2	22	17	10	7	23	21	51	11	2	15	59	72	14	4	3	9	11	12	36	42	87	16	21	27	9	10	25	66	47	15	15	Z
·————	1	2	5	10	3	3	5	13	4	2	29	5	14	9	7	14	30	28	9	4	2	3	12	14	17	19	22	84	63	13	8	10	8	19	32	57	55	1
··———	2	7	14	22	5	4	20	13	3	25	26	9	14	2	3	17	37	28	6	5	3	6	10	11	17	30	13	62	89	54	20	5	14	20	21	16	11	2
···——	3	3	8	21	5	4	32	6	12	2	23	6	13	5	2	5	37	19	9	7	6	4	16	6	22	25	12	18	64	86	31	23	41	16	17	8	10	3
····—	4	6	19	19	12	8	25	14	16	7	21	13	19	3	3	2	17	29	11	9	3	17	55	8	37	24	3	5	26	44	89	42	44	32	10	3	3	4
·····	5	8	45	15	14	2	45	4	67	7	14	4	41	2	0	4	13	7	9	27	2	14	45	7	45	10	10	14	10	30	69	90	42	24	10	6	5	5
—····	6	7	80	30	17	4	23	4	14	2	11	11	27	6	2	7	16	30	11	14	3	12	30	9	58	38	39	15	14	26	24	17	88	69	14	5	14	6
——···	7	6	33	22	14	5	25	6	4	6	24	13	32	7	6	7	36	39	12	6	2	3	13	9	30	30	50	22	29	18	15	12	61	85	70	20	13	7
———··	8	3	23	40	6	3	15	15	6	2	33	10	14	3	6	14	12	45	2	6	4	6	7	5	24	35	50	42	29	16	16	9	30	60	89	61	26	8
————·	9	3	14	23	3	1	6	14	5	2	30	6	7	16	11	10	31	32	5	6	7	6	3	8	11	21	24	57	39	9	12	4	11	42	56	91	78	9
—————	0	9	3	11	2	5	7	14	4	5	30	8	3	2	3	25	21	29	2	3	4	5	3	2	12	15	20	50	26	9	11	5	22	17	52	81	94	

An MDS of the Symmetrized Morse Code Data

Let us now study the symmetrized data by ordinal MDS.[1] In comparison with the color data examined above, we start here from a weaker position. Previously, we had a clear expectation about the MDS configuration and its dimensionality; here we make no attempt to predict anything. Hence, we must proceed in a purely descriptive way at the beginning.

Proceeding as Kruskal (1964a) did, we compute solutions in 1D through 5D, for which we obtain the Stress values shown graphically in Figure 4.5. The figure shows that the 1D solution has about .28 Stress. These values lie way below the expected Stress for random data reported in Figure 3.6, but that is always true for structured proximities. Adding one more dimension reduces Stress considerably to .18. By Kruskal's criteria (see Section 3.5), this would still be evaluated as a "poor" goodness-of-fit value. However, this simple norm does not take n, the number of points, into account, and what we have here is a relatively big data set compared to, say, the color data in Table 4.1. Fitting proximities for more objects to distances in an MDS space always requires a higher dimensionality if the data contain a certain amount of experimental error.

But how large is this error? We could take up the proposal of Spence and Graef (1974) and compare the observed Stress values to those obtained from simulating the Hefner model. This should allow us to determine both the true dimensionality and the error level. The observed Stress values are 0.35, 0.20, 0.14, 0.10, and 0.08 for $m = 1, \ldots, 5$, respectively. Their scree plot (Figure 4.5) shows no elbow. Turning to Figure 3.8, we note that the curves that most closely approximate the observed Stress values are the ones for an error level of 0.13. However, the Spence and Graef (1974) simulations do not clearly indicate what the true dimensionality of the MDS configuration is for these data.

Turning to interpretability, we first consider the 2D MDS configuration in Figure 4.6. Interpretation means to link some of the configuration's geometrical properties to known or assumed features of the represented objects. In the given case, we find that the points arrange themselves in a pattern that reflects the composition of the represented Morse signals, as shown in Figure 4.7. Following a suggestion by Wish (1967), we note that the 2D MDS space can be cut by the solid lines such that each region of the space contains signals of the same total duration. For example, this puts M (coded as 22), R (=121), D (=211), U (=112), and H (=1111) into the same equivalence class, because their signals all last 35/100 sec.

---

[1]The first MDS analysis of these data was done by Shepard (1963) and then by Kruskal (1964a) with the program M-D-SCAL (Kruskal & Carmone, 1969). M-D-SCAL has been replaced, in the meantime, by KYST (Kruskal, Young, & Seery, 1978). Most modern MDS programs (see Appendix A for an overview) usually lead to very similar solutions (Spence, 1972).

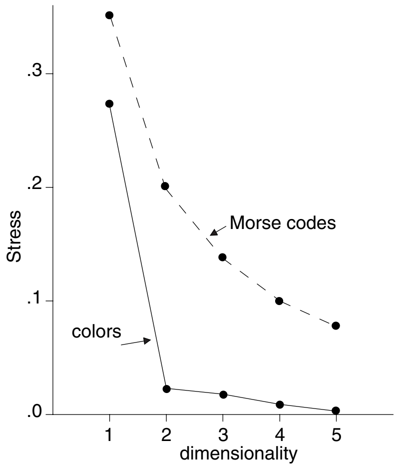

FIGURE 4.5. Scree plot (Stress vs. dimensionality) for MDS of color and Morse code data, respectively.

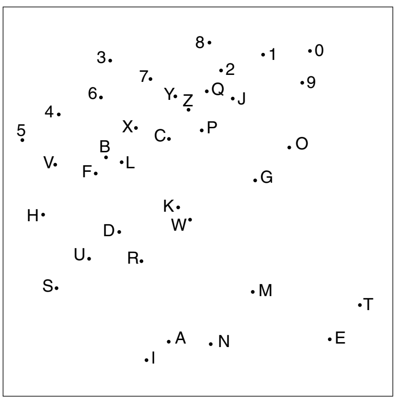

FIGURE 4.6. Ordinal MDS representation of Morse code data in Table 4.2.

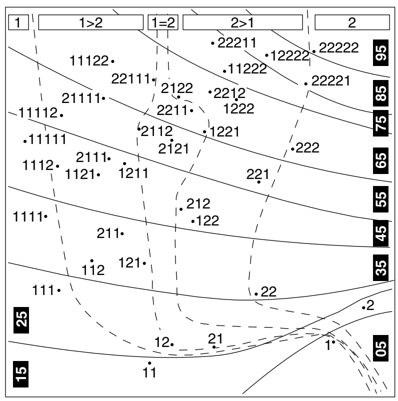

FIGURE 4.7. Morse code MDS configuration with two sets of partitioning lines: dashed lines split space into different signal types; solid lines differentiate signal lengths.

Technically, cutting a space into regions is called *partitioning* the space. Generally, partitioning a set means splitting it into subsets such that each element belongs to exactly one such subset. The resulting subsets are *exhaustive* and *disjoint*.

The configuration can also be partitioned in other ways by using other criteria. The dashed lines partition the space into regions that contain signals with only short (coded as 1) beeps, more short than long (coded as 2) beeps, a balanced number of short and long beeps, more long than short beeps, and long beeps only, respectively. The structure of this partitioning could be simplified—provided we are admitting some minor and one major misclassification of points—to a North–South slicing of the MDS plane into parallel stripes. The one major misclassification would result from point E. E, the Morse code that consists of one short beep only, seems to play a particular role. It is close to T, the other one-beep Morse code.

Without E, a typical *dimensional interpretation* of the MDS space would suggest itself: after a little rotation, the Y-axis could be interpreted as "duration", the X-axis as "kind of composition", ranging from signals consisting of short beeps only over signals with both short and long beeps to signals with long beeps only. Hence, at this stage, further research should first clarify the reliability of E's position. If E turns out to be reliable, we could possibly design a theory that explains the subjective similarity of Morse codes not by two independent dimensions but by two dimensions where the points' variance with respect to one dimension depends on the scale values on the other dimension, giving rise to a fan-like partitioning.

In any case, we see that the 2D MDS configuration can be interpreted in a simple but nontrivial way. Known properties of the signals, not just plausible posthoc insights, are used to explain the point scatter. The simplicity of the resulting geometric structure suggests, moreover, that we have found something real, not just an apparent structure in random data.

If we go on to higher-dimensional solutions, the points do not appear to reflect further systematic structure. Because no substantive hypothesis could be derived on the dimensionality of the MDS configuration, we may decide to give considerable weight to this simple interpretability of the solution over a formal precision-of-representation criterion such as Stress. This turns out to be a fruitful strategy in general. In any case, the data could be replicated, and then we would hope to find the same organizational patterns again. Without several such replications, we should be wary of making fine-grained interpretations.

4.3 Dimensions of Facial Expressions

There are many principles that can be used for interpreting an MDS configuration. What one always looks for is some way to organize the point scatter, to account for it or to "explain" it by a parsimonious but substantively meaningful generating function. The typical, often almost mechanical approach to this question in the literature has been the interpretation by dimensions. Dimensional interpretations assign substantive meaning to coordinate axes. We now examine a relatively refined example where a dimensional theory is given a priori.

Rating Facial Expressions on Simple Scales

Some of the early research on the psychology of facial expressions was occupied with the question of whether subjects could correctly identify the intended emotional message from a person's facial expression. It was found that misinterpretations were not random; the perceived emotion usually seemed "psychologically similar" (Woodworth, 1938) to the one actually expressed by the sender. Schlosberg and others then attempted to develop a theory of the differentiability of facial expressions, concluding that three perceptual "dimensions" were needed for a meaningful classification of facial expressions: pleasant–unpleasant (PU); attention–rejection (AR); and tension–sleep (TS). In different studies, it could be shown that subjects were able to classify facial expressions on these dimensions.

Engen, Levy, and Schlosberg (1958) published scale values, empirically arrived at, for the 48 photographs of the Lightfoot Series. This series shows the face of a woman acting out a series of different situations. Some of the situations and their coordinate values are given in Table 4.3. If these values are taken as Cartesian coordinates, distances between the different expressions can be computed and used to predict confusion rates. However, "... the particular three dimensions used by Schlosberg are not necessarily the only dimensions or the best dimensions for explaining confusion data .... There is the possibility that one or more of Schlosberg's scales, while understandable when made explicit to judges, are unimportant in uninstructed perception of facial expression; or conversely, that one or more important scales have been omitted .... [The experimenter] imposes particular dimensions of his own choosing and is arbitrarily forced to give them equal weight" (Abelson & Sermat, 1962, p. 546).

MDS of Facial Expressions and Internal Scales

MDS offers another way of testing the theory of three dimensions. We can ask the subjects to globally judge, without external criteria provided by the experimenter, the overall similarities of different facial expressions. The proximities are then mapped into MDS distances. The resulting configuration should be three-dimensional, with dimensions that correspond to the Schlosberg scales.

Abelson and Sermat (1962) asked 30 students to rate each pair of the 13 pictures described in Table 4.3 on a 9-point scale with respect to overall

TABLE 4.3. Scale values on three scales for faces of a woman acting different scenes (Engen et al., 1958); values are medians on 9-point scales.

	Scene	PU	AR	TS
1	Grief at death of mother	3.8	4.2	4.1
2	Savoring a Coke	5.9	5.4	4.8
3	Very pleasant surprise	8.8	7.8	7.1
4	Maternal love–baby in arms	7.0	5.9	4.0
5	Physical exhaustion	3.3	2.5	3.1
6	Something wrong with plane	3.5	6.1	6.8
7	Anger at seeing dog beaten	2.1	8.0	8.2
8	Pulling hard on seat of chair	6.7	4.2	6.6
9	Unexpectedly meets old boyfriend	7.4	6.8	5.9
10	Revulsion	2.9	3.0	5.1
11	Extreme pain	2.2	2.2	6.4
12	Knows plane will crash	1.1	8.6	8.9
13	Light sleep	4.1	1.3	1.0

dissimilarity. Dissimilarity was defined as “a difference in emotional expression or content.” For each subject, 78 proximities resulted, which were then rescaled over individuals by the method of successive intervals (Diederich, Messick, & Tucker, 1957). The means of these intervals were taken as the proximity data (Table 4.4).

We now analyze the data in Table 4.4 by ordinal MDS. The resulting Stress values for 1D up to 5D solutions are .24, .11, .06, .04, and .02, respectively. On purely formal grounds, we would probably decide that the 2D solution is reasonably accurate. However, because we are particularly interested in testing Schlosberg’s theory of three dimensions, we should also consider the 3D solution. To make things simpler, we first start with the 2D solution.

The point coordinates of the 2D solution (Figure 4.9) are shown in Table 4.5. One can check that the values in each column add up to zero. Geometrically, this means that the MDS configuration is *centered*; that is, its center of gravity lies at the origin of the coordinate axes. The coordinate vectors are also uncorrelated. This is so because the MDS configuration has been rotated to its principal axes orientation or, expressed differently, because the dimensions X and Y are the principal axes (see also Section 7.10) of this plane. Principal axes (PAs) are always uncorrelated.[2] The PAs can be found by locating an axis so that it accounts for as much of the points’ scattering as possible. That is, an axis is located such that it lies as close as possible to all points in the sense that the sum of squared distances of

[2] One can formulate the problem of finding PAs as finding that rotation of a given Cartesian dimension system that makes the point coordinates uncorrelated [see, for example, Strang (1976)].

TABLE 4.4. Proximities for faces from Table 4.3.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	–												
2	4.05	–											
3	8.25	2.54	–										
4	5.57	2.69	2.11	–									
5	1.15	2.67	8.98	3.78	–								
6	2.97	3.88	9.27	6.05	2.34	–							
7	4.34	8.53	11.87	9.78	7.12	1.36	–						
8	4.90	1.31	2.56	4.21	5.90	5.18	8.47	–					
9	6.25	1.88	0.74	0.45	4.77	5.45	10.20	2.63	–				
10	1.55	4.84	9.25	4.92	2.22	4.17	5.44	5.45	7.10	–			
11	1.68	5.81	7.92	5.42	4.34	4.72	4.31	3.79	6.58	1.98	–		
12	6.57	7.43	8.30	8.93	8.16	4.66	1.57	6.49	9.77	4.93	4.83	–	
13	3.93	4.51	8.47	3.48	1.60	4.89	9.18	6.05	6.55	4.12	3.51	12.65	–

TABLE 4.5. Coordinates for points in 2D MDS space.

Point/Picture	Dim 1 (X)	Dim 2 (Y)
1	−0.41	−0.46
2	0.54	0.14
3	1.22	0.75
4	0.97	−0.21
5	0.06	−0.72
6	−0.67	0.24
7	−1.34	0.45
8	0.48	0.62
9	1.05	0.27
10	−0.59	−0.69
11	−0.62	−0.31
12	−1.02	0.98
13	0.32	−1.04

the points from it is minimal. The second PA then is fixed automatically, because it must be perpendicular to the first axis.

Internal and External Scales

We now test whether the *external* scales of Table 4.3 account for the relative locations of the points. A crude first test is to correlate each of the columns of Table 4.5 (*internal scale*) with the columns in Table 4.3. Table 4.6, left panel, shows that there is a considerable correlation, $r = .94$, between the coordinates of the points on the X-axis and the values of the corresponding facial expressions on the PU scale. Similarly, the point coordinates on the Y-axis correlate highly with both the AR ($r = .86$) and the TS ($r = .87$) scales.

TABLE 4.6. Correlations between principal axes of 2D and 3D MDS solutions and Schlosberg scales in Table 4.3.

	2D MDS			3D MDS			
Scale	Dim 1	Dim 2	R^2	Dim 1	Dim 2	Dim 3	R^2
PU	.94	.21	.92	.93	.20	-.09	.91
AR	-.02	.86	.74	-.05	.83	-.34	.81
TS	-.38	.87	.90	-.37	.89	.06	.96

Yet, Schlosberg's theory does not claim that the principal axes should be of particular substantive importance. Maybe there are other dimensions that better satisfy the theory and, in particular, correlate higher with the scales of Table 4.3. This question can be answered as follows. Using multiple correlation, we can assess how well an optimal linear combination of the principal axes explains the scales. Because principal axes are uncorrelated, the squared multiple correlations are simply the sum of the squared bivariate correlations in Table 4.6. For example, for the PU scale on the one hand and the principal axes of the 2D solution, we find $R(PU.12) = .92^{1/2}$ from $R^2 = (0.94)^2 + (0.21)^2 = 0.92$. Thus, because the multiple correlation of the PU scale with the principal axes is higher than any bivariate correlation of PU with a given principal axis, there must exist an axis (i.e., another internal scale) in the MDS space that correlates even higher with PU than the X-axis. This is now investigated.

Optimally Fitting External Scales

In addition to correlating the points' coordinates on some internal scale with an external scale, we can also express their relationship geometrically. This is done by representing an external scale S by a directed line[3] Q located such that the point projections on it (Q-values or Q-coordinates) mirror as closely as possible the corresponding scale values of S. This can mean, for example, that the point projections on Q are spaced such that the ordinal Stress between the Q- and the S-values is minimal. Or, because we have treated the S scales above as interval scales, we could require that the intervals of the Q- and the S-values correspond most closely in their proportions. Thus, Q should be located such that, over all points i, $[s_i - (a + b \cdot q_i)]^2 = \min$, where q_i is the coordinate value of point i's projection on line Q. This looks like a linear regression problem, except that not only the weights a and b, but also the q_i values are unknowns. But any line Q is simply a linear combination of the coordinate vectors in

[3] A *directed line* is a line on which the direction from one end to the other has been indicated as positive, and the reverse direction as negative. The points on this line are ordered.

TABLE 4.7. Multiple regression problem to account for external PU scale by MDS coordinate vectors. Weights w_1, w_2 and additive constant a are to be chosen such that $\approx$ means "as nearly equal as possible." Optimal values are $g_1 = 2.679, g_2 = 0.816, a = 4.523$.

$$
\begin{bmatrix} 3.8 \\ 5.9 \\ 8.9 \\ 7.0 \\ 3.3 \\ 3.5 \\ 2.1 \\ 6.7 \\ 7.4 \\ 2.9 \\ 2.2 \\ 1.1 \\ 4.1 \end{bmatrix}
\approx g_1 \begin{bmatrix} -0.41 \\ 0.54 \\ 1.22 \\ 0.97 \\ 0.06 \\ -0.67 \\ -1.34 \\ 0.48 \\ 1.05 \\ -0.59 \\ -0.62 \\ -1.02 \\ 0.32 \end{bmatrix}
+ g_2 \begin{bmatrix} -0.46 \\ 0.14 \\ 0.75 \\ -0.21 \\ -0.72 \\ 0.24 \\ 0.45 \\ 0.62 \\ 0.27 \\ -0.69 \\ -0.31 \\ 0.98 \\ -1.04 \end{bmatrix}
+ a \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}
= \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \\ q_7 \\ q_8 \\ q_9 \\ q_{10} \\ q_{11} \\ q_{12} \\ q_{13} \end{bmatrix}
= \begin{bmatrix} 3.05 \\ 6.08 \\ 8.40 \\ 6.95 \\ 4.10 \\ 2.92 \\ 1.30 \\ 6.32 \\ 7.56 \\ 2.38 \\ 2.61 \\ 2.59 \\ 4.53 \end{bmatrix}
$$

Table 4.7. Hence, for each point i, it holds that $q_i = w_1 \cdot x_i + w_2 \cdot y_i$, where x_i and y_i are the coordinate values of point i on the given X- and Y-axes, respectively, and w_k is a weight.

Inserting this expression for q_i into the above loss function, we note that b can be pulled into the w_is so that the multiple regression problem in Table 4.7 emerges, where $g_i = b \cdot w_i$. Because X and Y are uncorrelated, the weights in Table 4.7 are simple regression weights. The additive constant a is simply the mean of the external scale. (One can eliminate a entirely by transforming the s_i-values into deviation scores.) The regression equation thus says that, given some point P such as point 1 with coordinates $(-0.41, -0.46)$, its corresponding q_1-value is $2.679 \cdot (-0.41) + 0.816 \cdot (-0.46) + 4.523 = 3.05$.

Overall, the resulting point coordinates on Q correlate with the external scale PU with .96, which checks with the $R^2 = .92$ from Table 4.6.

For the origin $O = (0.00, 0.00)$, we get $q_O = 0.00$, and so it is convenient to run Q through the origin O. For actually drawing Q in an MDS space, we have to find a second point on Q besides the origin. It can be shown that the regression weights g_i are the coordinates of such a point, provided Q runs through the origin O. Hence, we have two points that lie on Q, and this determines the line. In the given case, these points are $O = (0.00, 0.00)$ and $(2.679, 0.816)$.

A second possibility is locating the line Q on the basis of its angles to the coordinate axes. The direction cosine[4] of line Q with the ath coordinate

[4]The *direction cosine* of Q with the coordinate axis A_i is the cosine of the angle that rotates the positive end of Q onto the positive end of A_i.

TABLE 4.8. Coordinates of points $1, \ldots, 13$ of Fig. 4.8 projected onto the axes A_1, A_2, and A_3 of Fig. 4.8; r is the correlation of axis A_i with the PU values in Table 4.3.

	A_1	A_2	A_3
1	-0.564	0.136	0.526
2	0.802	-0.421	-0.368
3	1.748	-0.719	-1.133
4	1.479	-0.980	-0.338
5	0.165	-0.417	0.475
6	-1.029	0.722	0.170
7	-2.049	1.426	0.356
8	0.655	-0.121	-0.672
9	1.544	-0.810	-0.709
10	-0.811	0.185	0.774
11	-0.895	0.399	0.528
12	-1.635	1.412	-0.173
13	0.590	-0.811	0.564
$r =$	0.920	-0.780	-0.800

axis can be computed directly by the formula $\alpha_a = \cos^{-1}(g_a / \sum_{a=1}^{m} g_a^2)$, where g_a is the regression weight of the ath coordinate axis.

Because of the close relationship between regression weights and direction angles, we can conceive of the problem of representing an external scale by a line as a rotation problem: the task is to turn a line running through the origin such that the projections of the points on it correspond best to a given external scale. Figure 4.8 demonstrates this notion. A line or, rather, a directed axis is spun around the origin until it reaches an orientation where the points of the MDS configurations project on it so that these projections correlate maximally with the external scale. Three axes (A_1, A_2, and A_3) are shown graphically. The corresponding point projections are exhibited in Table 4.8. The table also shows the correlations of the projections with the scale values for the external PU scale from Table 4.3. One notes that A_1 has a high positive correlation with the PU scale, which indicates that the X-axis of the MDS solution can be interpreted as a continuum ranging from unpleasant to pleasant (see also Figure 4.9).

3D MDS of the Faces Data with Embedded External Scales

Because Schlosberg's theory is a theory of three dimensions, we also take a look at the 3D MDS solution. Figure 4.10 exhibits this configuration, together with the embedded external scales. Before going into further interpretations, we note that such a 3D configuration is not easy to look at, because what we see here is only a projection of this configuration onto a plane. The reader always has to mentally reconstruct the original configuration from this projection, which is often a difficult task. We note,

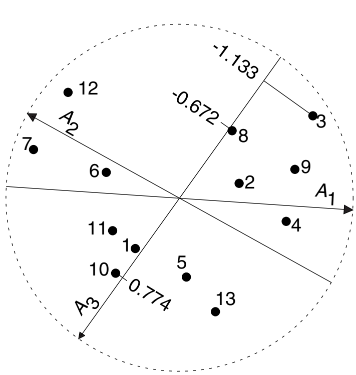

FIGURE 4.8. Embedding of an external scale into an MDS configuration (faces data) as a rotation problem.

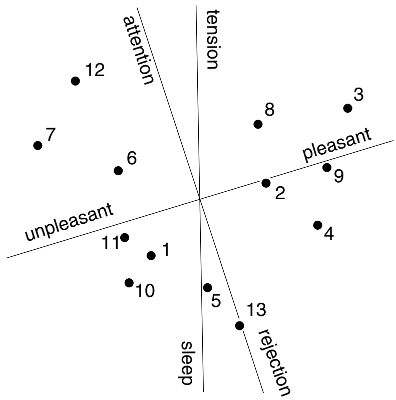

FIGURE 4.9. 2D MDS of faces data, with optimally fitted external scales.

for example, that it is almost impossible to see from Figure 4.10 how the embedded scales are oriented in the space.

One gets a clearer picture from the correlations of the embedded scales with the coordinate axes (Table 4.6). In addition, it is sometimes worthwhile to make use of features offered by the graphical environment of some MDS programs, in particular the possibility of rotating 3D configurations online in space. This allows one to inspect the configuration from different perspectives on the computer screen, which may suffice to understand the spatial relationships.

Figure 4.10, in any case, seems to suggest that the external scales PU and TS essentially correspond to Cartesian dimensions, whereas AR does not explain much additional variance. This is not surprising because r(TS, AR) = .75 in Table 4.3. Yet, there is quite a bit of scatter of the points in the third dimension. That this can only be partially explained by the external scales may be a consequence of the different psychology involved in generating the global similarity judgments and the ratings on the external scales. The given evidence is at least not contradictory to Schlosberg's theory of three dimensions.

4.4 General Principles of Interpreting MDS Solutions

The above MDS applications are chosen to show the reader some real-data examples, with substantive questions linked to them. The question of interpretation asked for connections of geometric properties of the MDS

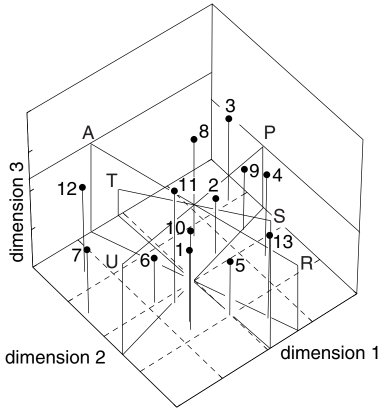

FIGURE 4.10. 3D MDS of faces data, with fitted external scales.

representation and substantive aspects of the represented objects. In the case of the color data, we found that the electromagnetic wavelengths of the colors were reflected in a corresponding array of the points in MDS space along a (curved) line. For the Morse code signals, we found that certain physical properties of the signals had a systematic relationship to various regions of the MDS space. The facial expression data led to an MDS configuration whose dimensions were strongly related to external scales for the same stimuli.

These examples illustrate the three most common principles used in interpreting MDS solutions. The color circle is an instance of a particular *manifold*, which is any set of points that form objects in space that are nearly "flat" in the neighborhood of any of their points ("locally" Euclidean). Most often, manifolds refer to points that form smooth curves or surfaces in space.

The regional interpretation of the Morse code data resulted from partitioning the space in multiple ways. The criteria used were different physical properties of the Morse code stimuli. In each case, the goal was to split the space such that each region would contain only points representing stimuli with equivalent properties on the partitioning criterion. Nothing else is required by this interpretational approach and, therefore, many different regional patterns may arise. Special cases are clusters—that is, very dense regions separated from each other by "empty space"—and dimensions. The latter partition the space into intervals, checkerboard patterns, box-like cells, and so on, depending on the dimensionality m. A regional interpretation is also possible for the color data: if we use wavelength as the physical property of the stimuli, each region contains but a single point, but coarser partitionings result from lumping together the stimuli into such classes as red, blue, yellow, and green.

Finally, the facial expression example illustrated the dimensional approach, the most common interpretation in practice. Note, however, that interpreting dimensions means that one is trying to link a very particular geometric feature to substantive features of the represented objects. One should not expect that this will always be successful.

These applications were, in a sense, confirmatory ones, because, in each case, there was at least an implicit expectation about certain properties of the MDS configuration. But even in a more exploratory context, interpreting MDS configurations complies with the same logic, except that some of the features of the stimuli one links to the MDS geometry are hypothesized or assumed. That is, looking at an MDS configuration and trying to make sense out of it simply means that one projects various forms of prior knowledge onto this space in order to explain the configuration. If this prior knowledge is solid, then exploratory MDS is also solid. Otherwise, one has to test the stability of such interpretations over replications.

In principle, *any* geometric property of an MDS solution that can be linked to substance is an interesting one. However, in the literature, certain standard approaches for interpretation are suggested, that is, particular geometric properties that one should consider. By far, the most popular approach is to look for meaningful directions or dimensions in the MDS space. Naturally, dimensions may not be related in any interesting way to the objects' substance, nor is any other feature of an MDS configuration.

4.5 Exercises

Exercise 4.1 In this exercise, we have a closer look at the choice of dimensionality for the color data of Ekman (1954) from Section 4.1.

(a) Compute MDS solutions for the data in Table 4.1 in 1, 2, 3, 4, 5, and 6 dimensions. Make a scree plot. What do you conclude with respect to the proper dimensionality of the MDS solution?

(b) Discuss a few criteria from Section 3.5 for choosing the proper dimensionality of the MDS solution.

Exercise 4.2 Figure 4.1 gives an MDS representation for the subjective similarity assessments of different colors. These colors are characterized by their electromagnetic wavelengths. Yellow corresponds to about 570 nm, green to 520 nm, blue to 480 nm, and violet to about 380–450 nm. Orange starts at about 600 nm and turns into red at the end of the visible spectrum (above 650 nm). For answering (b) and (c), you may want to consult an introductory psychology textbook.

(a) With this background, interpret the MDS configuration in terms of two meaningful color dimensions.

(b) What kind of MDS configuration could be expected if the color stimuli would vary not only in hue, but also in saturation?

(c) What color would you expect to lie at the center of the circle?

Exercise 4.3 Consider the facial expression data of Section 4.3.

(a) Compute the angle between the X-axis and the lines that best represent the scales PU, AR, and TS, respectively, of Table 4.3 in the MDS configuration of Table 4.5.

(b) The angles for AR and TS are similar. What does that mean in terms of the data?

(c) What substantive conclusions do you draw from (b)?

Exercise 4.4 Rosenberg and Kim (1975) studied the similarity of 15 kinship terms. College students sorted the terms on the basis of their similarity into groups. Each student generated a dissimilarity matrix where a pair of objects was coded as 1 if the objects were sorted in different groups and as 0 if the objects were sorted in the same group. The table below gives the percentage of how often terms were *not* grouped together over all students.

Kinship Term	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 Aunt	–														
2 Brother	79	–													
3 Cousin	53	67	–												
4 Daughter	59	62	74	–											
5 Father	73	38	77	57	–										
6 Granddaughter	57	75	74	46	79	–									
7 Grandfather	77	57	76	77	51	57	–								
8 Grandmother	55	80	78	54	70	32	29	–							
9 Grandson	79	51	72	72	54	29	31	57	–						
10 Mother	51	63	79	31	29	56	75	50	79	–					
11 Nephew	56	53	51	74	59	74	58	79	51	81	–				
12 Niece	32	76	53	52	81	51	79	58	74	60	27	–			
13 Sister	58	28	70	37	63	50	79	57	75	39	76	53	–		
14 Son	80	38	73	29	32	72	55	78	47	57	52	74	62	–	
15 Uncle	27	57	51	80	51	80	55	77	58	73	33	56	79	59	–

In addition, for each of the kinship terms, external scales can be set up for gender (1 = male, 2 = female, 9 = missing), generation (-2 = two back, -1 = one back, 0 = same generation, 1 = one ahead, 2 = two ahead), and degree (1 = first, 2 = second, etc.) of the kinship term. The table below presents these external scales.

Kinship Term		Gender	Generation	Degree
1	Aunt	2	-1	3
2	Brother	1	0	2
3	Cousin	9	0	4
4	Daughter	2	1	1
5	Father	1	-1	1
6	Granddaughter	2	2	2
7	Grandfather	1	-2	2
8	Grandmother	2	-2	2
9	Grandson	1	2	2
10	Mother	2	-1	1
11	Nephew	1	1	3
12	Niece	2	1	3
13	Sister	2	0	2
14	Son	1	1	1
15	Uncle	1	-1	3

(a) Do an ordinal multidimensional scaling analysis in two dimensions. Interpret the solution.

(b) Inspect the Shepard diagram or the transformation and residual diagrams. Are all proximities properly fitted?

(c) Compute the correlations between the dimensions and the external scales generation and degree, respectively. Use a multiple regression program to find optimal weights g_1 and g_2 to predict each external scale out of the two dimensions. Plot the two external scales in the solution. How can you interpret the solution in terms of generation and degree?

(d) Suppose that we would also like to represent gender in the MDS solution. Explain how this could be done. Elaborate your solution in the plot.

Exercise 4.5 Wolford and Hollingsworth (1974) were interested in the confusions made when a person attempts to identify letters of the alphabet viewed for some milliseconds only. A confusion matrix was constructed that shows the frequency with which each stimulus letter was mistakenly called something else. A section of this matrix is shown in the table below.

Letter	C	D	G	H	M	N	Q	W
C	–							
D	5	–						
G	12	2	–					
H	2	4	3	–				
M	2	3	2	19	–			
N	2	4	1	18	16	–		
Q	9	20	9	1	2	8	–	
W	1	5	2	5	18	13	4	–

(a) Are these data similarity or dissimilarity measures?

(b) Use MDS to show their structure.

(c) Interpret the MDS solution in terms of regions. What do you conclude with respect to letter confusion? (Hint: Letter confusion may be based, e.g., on visual features or on the similarity of sounds.)

Exercise 4.6 Consider the data on the subjective similarity of different countries in Table 1.3. The table below supplements these data by two external scales. The first scale consists of rankings on "economic development" that one particular student could have assigned to these countries in the 1960s. The second scale shows the population of these countries in about 1965.

Country	No.	Economic Development	Population (ca. 1965)
Brazil	1	3	87
Congo	2	1	17
Cuba	3	3	8
Egypt	4	3	30
France	5	8	51
India	6	3	500
Israel	7	7	3
Japan	8	9	100
China	9	4	750
USSR	10	7	235
U.S.A.	11	10	201
Yugoslavia	12	6	20

(a) Find the coordinates of a two-dimensional ordinal MDS representation of the data in in Table 1.3.

(b) Fit the external scales into this MDS space by linear regression. Plot the embedded scales as directed lines.

(c) Interpret the MDS solution in terms of the external scales, if possible. Discuss how successful these two scales are in explaining the MDS configuration.