Papeles del Psicólogo, 2010. Vol. 31(1), pp. 46-56
http://www.cop.es/papeles

MULTIDIMENSIONAL SCALING: CONCEPT AND APPLICATIONS

Constantino Arce, Cristina de Francisco and Iria Arce
University of Santiago de Compostela

The present article offers a conceptual, and at the same time operative, vision of the concept of multidimensional scaling. In the manner in which it is presented, the aim is first of all to help interested psychologists understand the multidimensional scaling model, using a number of simple, intuitive examples; and, secondly, to help them acquire the competence required for solving different multidimensional scaling problems through the use of specific software. Mathematical formulae and methods will be downloaded in the process, though at no point shall we renounce the methodological rigour the subject demands.
Key words: *Scaling, Proximity data, Preference data, Dimensionality reduction.*

A través del presente artículo se ofrece una visión conceptual, a la vez que operativa, del concepto de escalamiento multidimensional. En la forma de presentación se busca, en primer lugar, que los psicólogos interesados comprendan lo que es el modelo de escalamiento multidimensional a través de varios ejemplos muy sencillos e intuitivos y, en segundo lugar, adquieran competencias que le permitan resolver distintos problemas de escalamiento multidimensional con el uso de software específico. Se pretende igualmente descargar la presentación de fórmulas y métodos matemáticos sin renunciar por ello al rigor metodológico que el tema requiere.
Palabras clave: *Escalamiento de objetos, Escalamiento de sujetos, Datos de proximidad, Datos de preferencia, Reducción de la dimensionalidad.*

Multidimensional scaling, in its most basic formulation, sets out to represent a set of objects in a low-dimensionality space. The word 'object' is quite generic, and actually refers to any entity one might wish to scale. Another, equivalent term in Psychology is stimulus. The number of dimensions, usually small (two, three, four), is decided by the researcher for substantive reasons, though it can also be set on the basis of statistical criteria. The models and construction methods of unidimensional scaling, which were developed in the first half of the 20th century (notably by Thurstone, Likert, Guttman or Coombs), constitute the antecedents of the most modern models and methods of multidimensional scaling; indeed, in many cases the latter can be considered as generalizations of the former.

The first author to develop a model and a method of multidimensional scaling was Torgerson (1958). Today his model is known as the classical metric model. The name metric derives from the measurement scale that is assumed, or required, for the data, which is based on intervals, in Stevens' hierarchy. A few years later, Shepard (1962) and Kruskal (1964a, 1964b) proposed a model that permits a descent in the measurement scale to the ordinal level. This model is referred to as classical non-metric. Carroll and Chang (1970) were responsible for a significant advance with their proposal of a model that permitted the derivation of, in addition to the object space, a subject space in which to represent the weight or weighting given by each subject to each dimension of the object space. The Carroll and Chang model, known as INDSCAL, is of considerable psychological interest, since it permits or takes into account individual differences in the perception of the object space. There is a common object space, shared by all the subjects, but it allows for differences between some individuals and others in the perception of that configuration.

There are specific computer programs for each one of the models mentioned above, but today it is possible to solve multiple multidimensional scaling problems with a single computer program, such as PROXSCAL or ALSCAL, which involve the implementation of numerous models and form part of the widely used SPSS statistical package.

One of the features that most clearly distinguishes multidimensional scaling from other statistical data analysis models is the input matrix. In Psychology we are accus-

Correspondence: Constantino Arce, *Facultad de Psicología, Universidad de Santiago de Compostela, 15.782 Santiago de Compostela. España. E-mail: constantino.arce @usc.es*

tomed to using a rectangular data matrix X with n subjects in the rows and p variables in the columns, where an element x_{ij} represents the measure obtained for a subject i in a variable j. In its most typical form, the input matrix for the multidimensional scaling is a square data matrix of order p with the same entity represented in the rows and the columns: the objects we are trying to represent in the multidimensional space. An element in this matrix (delta) represents the distance or dissimilarity between two objects i and j. What we have in the matrix is actually a matrix of distances or dissimilarities between all the pairs of objects.

The difference between distance (geometric concept) and dissimilarity (psychological concept) resides in the fact that the former, being a mathematical concept, does not contain error, whilst the latter, being a psychological, perceptual or subjective concept, does indeed contain error. Dissimilarities are, in reality, distances that contain error or distances distorted by the perceptual mechanisms of human beings. The models and methods of multidimensional scaling can solve both types of problem, with error and without error in the input data. In Psychology it is more customary to work with data that contain error, and multidimensional scaling models can provide an approach to this problem.

DERIVATION OF A CONFIGURATION OF POINTS FROM A DISTANCE MATRIX

Table 1 shows the kilometric distance matrix between 7 Spanish cities: A Coruña, Bilbao, Barcelona, Cáceres, Madrid, Sevilla and Valencia.

Our proposal was, on the basis of this matrix, to draw up a map of Spain – that is, to obtain a spatial representation of the 7 cities in a plane, in which one of the axes would be the direction north-south and the other axis would be east-west. To this end we used the PROXSCAL procedure, implemented in SPSS.

The result can be seen in Figure 1.

Given that the map of Spain is well known, we can rate subjectively the extent to which the map derived by the program equates to the true map. It can be said that the map obtained is quite good, though not perfect. In Psychology research it is usual to work with configurations without an objective counterpart known in advance. Therefore, when the program produces a solution it becomes very important to have an indicator, or even several (the more the better), of the degree to which the configuration derived by the program fits the ideal (unknown) one. All multidimensional scaling programs provide users with goodness-of-fit indices so that they can rate how "good" the solution obtained by the program is for their problem.

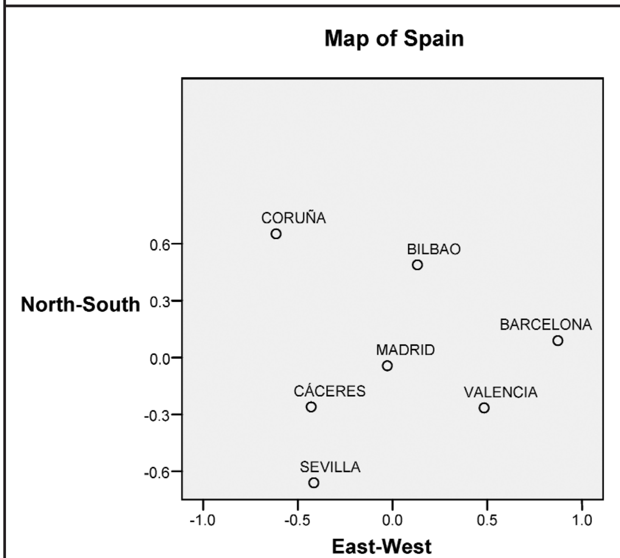
The goodness-of-fit indices offered by PROXSCAL for the map of Spain are shown in Table 2.

There are two types of indicator: those for which zero represents a perfect fit (of this first type would be the indicators Normalized Raw Stress, Stress-I, Stress-II and S-Stress); and those for which the perfect fit is represented by 1 (of this second type would be the last two in the table – Dispersion Accounted For [D.A.F.] and Tucker's congruence coefficient). Observing the values of any of these indicators one reaches the same conclusion: that the model's fit in this case is good or very good. This is

TABLE 1
KILOMETRIC DISTANCES BETWEEN 7 SPANISH CITIES

	A Coruña	Barcelona	Bilbao	Cáceres	Madrid	Sevilla	Valencia
A Coruña	0						
Barcelona	1050	0					
Bilbao	542	567	0				
Cáceres	617	895	591	0			
Madrid	586	600	379	294	0		
Sevilla	857	971	847	256	507	0	
Valencia	937	341	569	615	352	637	0

FIGURE 1
MAP OF SPAIN DERIVED BY PROXSCAL



because the degree of error in the data (distances) is very small. The distances used as input data were by road. If we were to use distances "as the crow flies" the fit would be perfect. The first four goodness-of-fit indices should equal 0 and the last two should equal 1.

PERCEPTION OF MEANS OF TRANSPORT

Arce (1993) set out to obtain a perceptual map of the means of transport used by the citizens of Santiago de Compostela. To this end, he drew up a list of all the means of transport (public and private) that could be available to them, made up all the pairs possible and asked a sample of citizens to judge the dissimilarity for each pair of means of transport.

Nine means of transport were studied: aeroplane, train, inter-city bus, city bus, taxi, private car, motorcycle, moped and bicycle. With nine objects or stimuli (in this case, means of transport) 36 pairs can be formed. The formula used to arrive at the number of pairs was n(n-1)/2, where n is the number of objects or stimuli. Substituting, in this case, where n = 9, we are left with 9(9-1)/2 = 36. Table 3 shows the 36 pairs formed in the study following the standard rotation method, which is very useful because the data (dissimilarities) are already in order in the form in which they will subsequently be entered in the input matrix. The method follows the sequence (1,2), (1,3) … (1,9), (2,3), (2,4) … (2,9), (3,4) (3,5) … (3,9) … (8,9).

For forming the number of pairs, we assumed the symmetry of the dissimilarity judgements, meaning that for a given pair (e.g., aeroplane/train) it is assumed that the judgement emitted by a subject will be the same if the pair is presented in the order aeroplane/train as if it were presented in the order train/aeroplane. With some rare exceptions, this assumption is customary in Psychology research.

Once the list with all the pairs to be judged by the subjects has been drawn up, a response scale must be designed for them to indicate their dissimilarity judgements on the objects or stimuli in each pair. The research in question employed a nine-point scale, in which 1 indicated that the means of transport included in each pair were very similar, and 9 indicated that they were very different. By way of example:

Aeroplane/train

Very similar	Moderately similar	Very different
1 2	3 4 5 6	7 8 9

Subjects used this scale to judge the dissimilarity in the 36 pairs formed in the study.

Figure 2 shows the perceptual map of the means of transport for one subject from the sample. In the configuration of points obtained we now have an additional problem with respect to the configuration in Figure 1. In the problem of distances between cities we knew the meaning of the axes: one was for the north-south direction and the other was for east-west. But what is the meaning of the axes of the configuration we have now obtained? The subject, in making his or her judgements, has probably used different axes or dimensions to assess the dissimilarity between the means of transport. For example, it may be that in judging the dissimilarity for a given pair he/she has concentrated on the safety of the

TABLE 2
GOODNESS-OF-FIT INDICES OFFERED BY PROXSCAL

Normalized Raw Stress	.00055
Stress-I	.02349
Stress-II	.06824
S-Stress	.00117
Dispersion Accounted For (D.A.F.)	.99945
Tucker's Congruence Coefficient	.99972

TABLE 3
LIST OF PAIRS OF MEANS OF TRANSPORT, ORDERED BY STANDARD ROTATION

1. Aeroplane/train	19. Inter-city bus/motorcycle
2. Aeroplane/inter-city bus	20. Inter-city bus/moped
3. Aeroplane/city bus	21. Inter-city bus/bicycle
4. Aeroplane/taxi	22. City bus/taxi
5. Aeroplane/private car	23. City bus/private car
6. Aeroplane/ motorcycle	24. City bus/ motorcycle
7. Aeroplane/moped	25. City bus/moped
8. Aeroplane/bicycle	26. City bus/bicycle
9. Train/inter-city bus	27. Taxi/private car
10. Train/city bus	28. Taxi/motorcycle
11. Train/taxi	29. Taxi/moped
12. Train/private car	30. Taxi/bicycle
13. Train/motorcycle	31. Private car/motorcycle
14. Train/moped	32. Private car/moped
15. Train/bicycle	33. Private car/bicycle
16. Inter-city bus/city bus	34. Motorcycle/moped
17. Inter-city bus/taxi	35. Motorcycle/bicycle
18. Inter-city bus/private car	36. Moped/bicycle

means of transport, while for another pair he/she may have focused on their social prestige, and so on. By means of multidimensional scaling we seek to obtain a configuration of points, but also to ascertain the meaning of each axis or dimension of that configuration. There are various ways of approaching this issue, but the most reliable involves gathering more data. In fact, in the above-mentioned research, in addition to asking subjects for their dissimilarity judgements, they were asked to rate each of the means of transport with respect to a series of properties including safety, stability, resistance, strength, weight, attractiveness, prestige, punctuality, social status and comfort. Subsequently, it was determined whether there was any type of relationship between any of these properties and the position of the means of transport on each one of the dimensions derived. In a first, exploratory phase we tried solutions with 2, 3 and 4 dimensions. The solution that produced the best meaning was that of 3 dimensions. Multiple regression analyses, in which we took as dependent variable a given property of the means of transport and as independent variables the coordinates of the means of transport derived with the multidimensional scaling programs, showed that dimension 1 (horizontal axis) represented the perceived safety of the means of transport, dimension 2 (vertical axis) referred to their attractiveness and dimension 3 (depth axis) represented their social prestige. In Figure 2, which shows the first two dimensions of the three-dimensional solution, the means of transport situated on the right (train, inter-city bus, city bus, etc.) are perceived as safer, and those on the left (bicycle, motorcycle, moped) as more unsafe. Likewise, the means of transport situated higher up (aeroplane, taxi, car, motorcycle) are perceived as more attractive, and those situated lower down as less attractive (train, city-bus, inter-city bus, bicycle, moped).

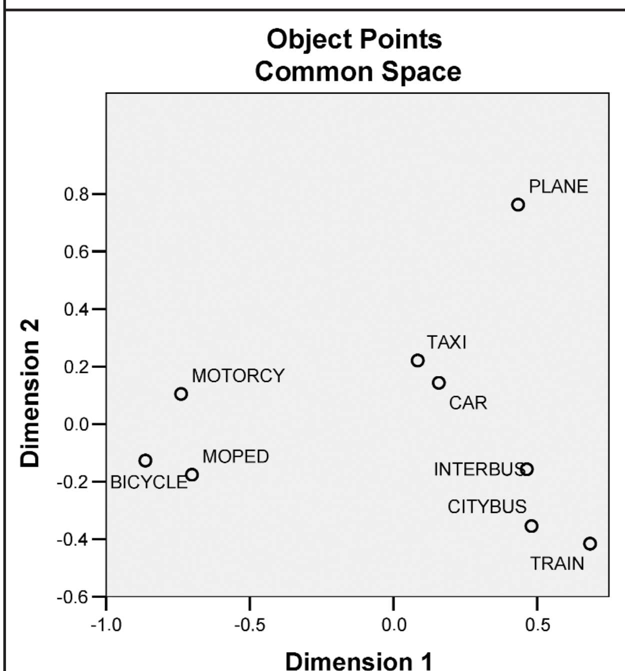
THE CASE OF MORE THAN ONE INPUT MATRIX: THE INDSCAL MODEL

The examples used so far involved one input matrix. In the first problem it was the matrix of kilometric distances between the seven Spanish cities; in the second problem it was the matrix of dissimilarities between the means of transport for a subject in the sample. In the first problem, indeed, there was no other possibility, because there is just one distance matrix, but the second problem involves multiple subjects, and we would have liked to enter the dissimilarity matrix of each subject. In fact, in the original research this is what was done. Today, any multidimensional scaling program will permit one to obtain a common object space shared by a sample of subjects or another data source.

Among the models that deal with the issue of multiple input matrices there is one that merits particular attention, because it has properties that can be very interesting from the psychological point of view: Carroll and Chang's (1970) INDSCAL model. This model permits us to obtain two spaces: the object space, common to all subjects in the sample, and the subject space. The novel aspect of the model is actually this latter space. Represented in the subject space is the weight, weighting or importance given by each subject to each one of the dimensions of the configuration of objects. Thus, the subjects share the same object space but the model permits each of them to perceive that space differently; in sum, it allows individual differences between the subjects.

Arce (1994) asked two subjects to rate the dissimilarity between 7 makes of car: Ferrari, Porsche, BMW, Mercedes, Renault, Seat and Opel. He obtained two dissimilarity matrices and used both for multidimensional

FIGURE 2
PERCEPTUAL MAP OF THE MEANS OF TRANSPORT
(DIMENSION 1 VERSUS DIMENSION 2)



scaling. The results revealed that the first perceptual dimension was the "sporty" features of the brand and the second referred to comfort. Figure 3 shows the common object space shared by the two subjects. The cars situated more to the right (dimension 1) are perceived as more sporty than those situated on the left, and those situated lower down (dimension 2) are considered more comfortable than those situated higher up.

FIGURE 3
OBJECT SPACE

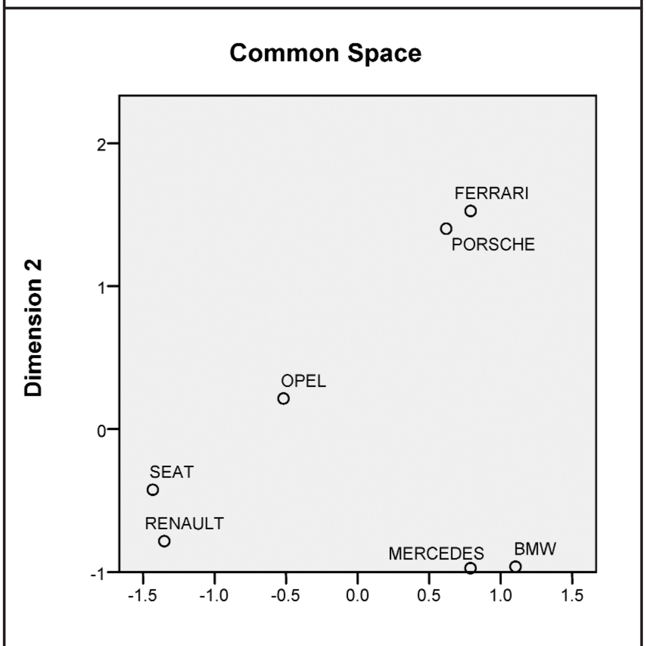


FIGURE 4
SUBJECT SPACE

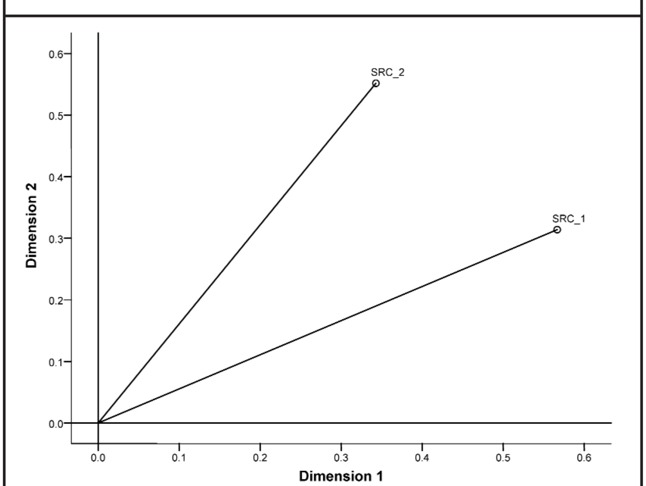


Figure 4 shows the subject space. Whilst the object space is common for the two objects, the subject space tells us that subject 1 (SRC_1 on the graph) gives more importance to dimension 1, the sporty features of the car, whilst subject 2 (SRC_2 on the graph) gives more importance to the comfort of the cars. In contrast to the case of the object space, where each object is represented by a point, in the subject space the subject is represented by a vector (a line). The nearer this vector is to a dimension, the greater the importance given by the subject to that dimension, and the further away it is, the less the importance attributed to it. Indeed, it can be seen in the graph that subject 1 is closer to dimension 1 (sporty design), thus giving more importance to this dimension, whilst subject 2 is closer to dimension 2 (vertical axis), indicating that it is this dimension (comfort of the cars) which carries more weight in his judgements on car brands.

MULTIDIMENSIONAL SCALING WITH PREFERENCE DATA

Although multidimensional scaling, in its most typical form, starts out from a matrix of dissimilarities among objects, some models and methods have been developed which permit the multidimensional scaling of objects on the basis of preference data (e.g., Bennett & Hays, 1960; Carroll, 1980; Tucker, 1960). If we have n objects that we wish to scale, we simply ask the subject to put them in order of preference, assigning the number 1 to the most preferred object, 2 to the second most preferred, and so on until the last object, to which the number n must be assigned. The advantage of these data is that they are much more conveniently obtained than dissimilarity data. The task tends to be much simpler for both subject and researcher. The preference data are subsequently ordered within a rectangular matrix, with subjects in the rows and objects in the columns. Each row is a subject, and an element of the row represents the preference which that subject has given to a particular object.

By way of an example, let us suppose that we were interested in obtaining a perceptual map of the sports and physical activities available to citizens in their leisure time. To make the example manageable we can choose 8 sports or physical activities and 16 subjects, who are asked to indicate their preferences by marking the sport or physical activity they most prefer with a 1, their second preference with a 2, and so on until their least preferred sport or physical activity, which they must mark with an 8.

The sports or physical activities selected in the example are: football, basketball, tennis, athletics, walking, swimming, cycling and running.

The preferences indicated by the subjects are shown in Table 4.

Figure 5 shows the perceptual map of the sports and physical activities rated by the subjects in the sample. To interpret the meaning of the dimensions we look, first of all, at the properties of the sports or physical activities placed at each extreme of the dimensions. They will probably have some property reflecting a contrast that can help us to discover the meaning of the respective dimension. Thus, we can observe that in dimension 1 (horizontal), situated on the right are the non-competitive physical activities (walking, swimming, running and cycling), and on the left are the competitive sports (football, athletics, tennis and basketball). Dimension 1 could be interpreted, therefore, as the competitiveness of the sports or physical activities. Similarly, if we look at the positioning of the sports and physical activities in dimension 2 (vertical axis), we can appreciate that toward the top are the sports or physical activities of an individual nature (athletics, running, tennis, etc.) and in the lower part are the team sports (football, basketball), so that this second dimension can be interpreted as the type of sport or activity: individual versus team-based.

SOLUTION OF MULTIDIMENSIONAL SCALING PROBLEMS WITH SPSS

Up to now, although we have referred to the process of solving scaling problems by means of specific software, we have not gone into much detail. We shall now illustrate how we solved the problem of kilometric distances

FIGURE 5
PERCEPTUAL MAP DERIVED BY PROXSCAL FOR THE SPORTS AND PHYSICAL ACTIVITIES BASED ON PREFERENCE DATA

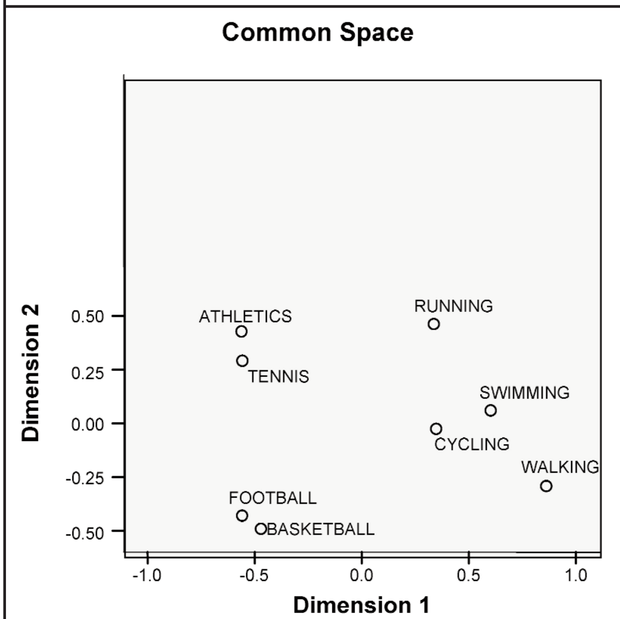


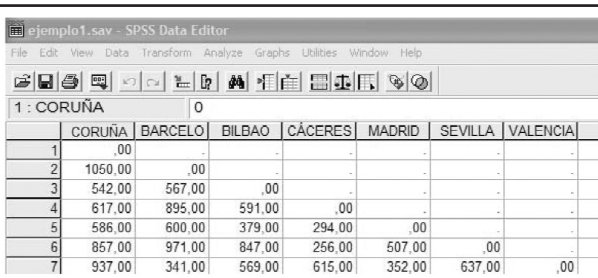
TABLE 4
SUBJECTS' PREFERENCES

Subject	Football	Basketball	Tennis	Athletics	Walking	Swimming	Cycling	Running
1	8	7	6	5	1	4	3	2
2	7	8	5	6	2	3	4	1
3	8	7	6	5	1	3	2	4
4	7	8	5	6	2	4	3	1
5	6	5	7	8	1	2	3	4
6	5	6	7	8	2	3	4	3
7	6	5	7	8	2	1	3	4
8	5	6	8	7	3	2	4	3
9	1	2	3	4	5	6	7	8
10	2	1	4	3	5	7	6	8
11	1	2	4	3	8	7	6	5
12	2	1	3	4	8	6	7	5
13	3	4	1	2	8	6	7	5
14	4	3	1	2	8	7	6	5
15	4	3	2	1	7	8	5	6
16	3	4	2	1	8	6	7	5

with the PROXSCAL procedure implemented in SPSS. At the same time, in some cases, we shall indicate the differences in the decision-making between this problem and the other three we have also solved.

Step 1. Create the data file in SPSS with the kilometric distances between the seven Spanish cities.
It should look like the one in Figure 6. Given that the input matrix is square, the rows in the matrix have the same meaning as the columns. Row 1 is A Coruña, row 2 Barcelona, and so on until row 7, which is Valencia. The fact that the names of the cities appear in the columns and not in the rows is because the SPSS system allows the columns of the data matrix to be labelled (using the Variable View tab), while it does not allow the naming of rows.

FIGURE 6
INPUT FILE FOR SPSS WITH KILOMETRIC DISTANCES BETWEEN THE SEVEN CITIES*


ejemplo1.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

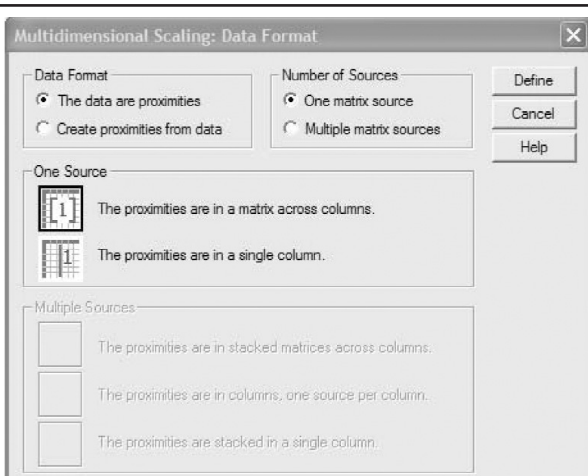
1 : CORUÑA 0

	CORUÑA	BARCELO	BILBAO	CÁCERES	MADRID	SEVILLA	VALENCIA
1	,00	.	.	.	.	.	.
2	1050,00	,00	.	.	.	.	.
3	542,00	567,00	,00	.	.	.	.
4	617,00	895,00	591,00	,00	.	.	.
5	586,00	600,00	379,00	294,00	,00	.	.
6	857,00	971,00	847,00	256,00	507,00	,00	.
7	937,00	341,00	569,00	615,00	352,00	637,00	,00

*Note to Fig. 6: the decimals (all zero) are denoted by commas rather than points, since the data was entered by the authors using the Spanish format.

*Note to Fig. 6: the decimals (all zero) are denoted by commas rather than points, since the data was entered by the authors using the Spanish format.

FIGURE 7
SELECTION OF DATA FORMAT AND NUMBER OF INPUT MATRICES IN THE PROXSCAL PROCEDURE



Step 2. Select the procedure we wish to run:
Analyze/Scale/Multidimensional Scaling (PROXSCAL)

Step 3. Data Format
The PROXSCAL procedure permits two types of input data:
(a) proximity data (square matrix)
(b) profile data (rectangular matrix)
Kilometric distances, like dissimilarities between objects, are proximity data. We therefore select the option which indicates to the program that the data are proximities (see Figure 7).
If we had a rectangular input matrix, such as in the case of the preferences in the example of sports or physical activities, then we would have to select the option that requests the program to create proximities from data.

Step 4. Number of input matrices
The PROXSCAL procedure permits one or more than one input matrix. The number of matrices we have in the problem is indicated in the box called Number of Sources. Since in this case we have only one matrix, we choose the option "One matrix source" (see Figure 7). In the third problem we have solved here, that of the makes of car, we chose the option "Multiple matrix sources", as we have two matrices, one per subject. In the input file the matrices are situated one below the other, the same format being maintained in all of them.

Step 5. Click on the Define button (see Figure 7)

Step 6. Select the objects we wish to scale (in this case, cities)
For this we select the 7 cities in the box on the left in Figure 8 and move them to the box called Proximities by clicking on the arrow between the two.

Step 7. Selection of the Model (click on the button "Model")
Since we have just one input matrix of proximities (distances), the procedure does not allow us to select the scaling model (see Figure 9), which will any case be similar to the classical model.

If we had more than one input matrix of proximities we would indeed be able to select the scaling model. The most usual would be the model with replication (called Identity model in the dialogue box) and the INDSCAL model (called Weighted Euclidean model in the dialogue box). In the replication model the subjects are considered as replications of one another, which means that the differences there may be between them are attributed to random factors. The INDSCAL model, on the other hand, permits individual differences. In the problem of the makes of car, we chose this model.

Step 8. More decisions about the format and nature of the data

Up to now, the program knows that we have an input matrix of proximities (square), but it still requires us, under the "Shape" heading (see Figure 9), to specify whether we have the information in the lower triangle, the upper triangle, or the full matrix. Since the distance matrix is symmetrical, we opted to have the information only in the bottom half (lower triangular matrix). The complete matrix option is only used when the input matrix is asymmetrical.

Multidimensional scaling programs project dissimilarities as distances in space. The greater the dissimilarity between objects, the greater will be the distance in the multidimensional space. But the input data can be dissimilarities or similarities. If they are similarities, the relationship with the distances will be inverse: the greater the similarity between two objects in the empirical world, the smaller will be the distance between them in the space. This is, then, an important specification that the user must make on employing the program. In our case, in the Proximities box, we must choose Dissimilarities. Distances are conceived as dissimilarities.

Under the "Proximity Transformations" heading (see Figure 9), the procedure permits the user to choose the measurement scale for the input data. If we choose ratio or intervals, the model will be metric, and if we choose ordinal it will be nonmetric. In Psychology research the "ratio" measurement level is rarely selected; the most common are intervals and ordinal. In the problem here, however, given that the data are distances and not subjects' judgements, we chose the highest measurement level (ratio).

In the sports problem, in which we used preference data, we specified that the measurement level was ordinal. In this case the program permits us to make still one more specification, under the heading "Untie tied observations". This is a highly technical decision. If we select it, the program will assume that the measurement process is continuous, and if we do not, that it is discrete. The decision has repercussions only in those cases in which there are ties. By default, the program assumes that the measurement process is discrete and respects ties in the data. If we consider that the measurement process is continuous, we must tell the program to untie the tied observations. In our sports problem we tested both options and noticed no differences in the solutions derived by the program. In fact, this is what happens in the majority of cas-

FIGURE 8
SELECTION OF OBJECTS WE WISH TO SCALE (IN THIS CASE, CITIES)

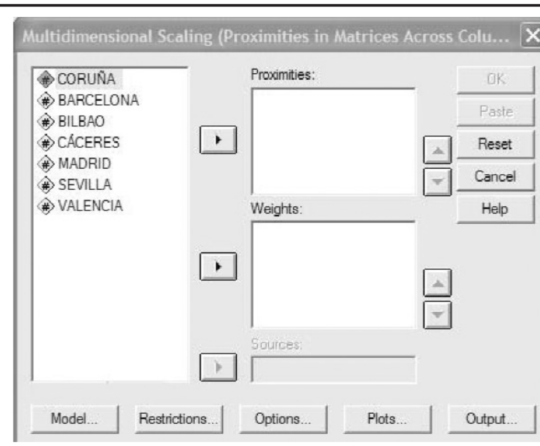
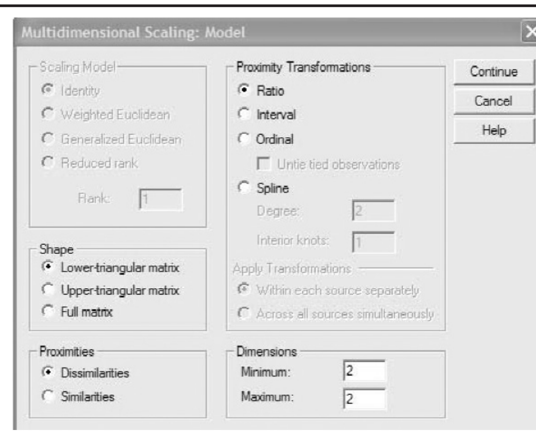


FIGURE 9
SELECTION OF THE SCALING MODEL



es; these decisions are important at a mathematical level, but not at a substantive one.

Step 9. Number of dimensions

If we have a clear starting hypothesis we can select a fixed number of dimensions, and if we do not have such a hypothesis, the best approach is to try to obtain different solutions and subsequently choose the solution with the number of dimensions that is most interpretable from a substantive point of view. In our example of the distances, the number of dimensions we chose, given that it was a map, was two (see Figure 9).

In the example of the means of transport, as can often occur in Psychology research of an exploratory nature, we did not have such a clear hypothesis with regard to the meaning of the dimensions we could obtain. Consequently, we tried to obtain solutions in two, three and four dimensions. We then attempted to seek a meaning for them a posteriori. Having found that three were interpretable, this was the solution we chose.

The interpretation of the dimensions can be made by the researcher, attempting to analyze first of all the properties of the objects occupying the most extreme positions in the dimension. When the procedure separates the objects considerably, it tends to be because they have some opposing properties which, if identified, can help us to name the dimension. We used this procedure in the sports example. Nevertheless, this interpretation based on the opinion of an expert (the researcher) may be called into question by other researchers (or experts). The ideal approach is to proceed as we did in the means of transport problem, where in addition to asking the subjects for dissimilarity judgements, we asked them to rate each type of transport on a series of bipolar scales that represented hypothetical properties of the means of transport. Subsequently, using statistical methods of correlation and regression, it was possible to offer evidence about the true meaning of each of the dimensions retained.

Step 10. Restrictions and Options

Both the Restrictions and the Options buttons (see Figure 8) permit the user to make decisions at a highly advanced level. In practice, it is customary to take the default options implemented by the procedure. In Restrictions, by default, the program assumes that it must estimate all the coordinates of the objects (without restrictions). Sometimes, exceptionally, the coordinates are known, and all that is sought is to project new objects onto an already defined space. In such a case it would be necessary to provide the program with the coordinates, which it would read from a file we indicate to it.

Likewise, in Options, by default, the program takes a particular initial configuration (simplex), which it permits us to change for other alternatives (e.g., Torgerson). It also allows us to change the criteria for attaining convergence and the number of iterations made by the algorithm. Rarely can better results be obtained if the program's default options are changed.

Step 11. Decision-making about output

The Plots and Output buttons of the program (see Figure 8) permit users to choose what they want to appear in the output. The possibilities range from a very simple output, with the substantial elements, to a highly elaborate output with all types of technical detail. It is advisable, initially, to obtain a simple output, and subsequently, if necessary, to obtain new outputs with more informative elements. By default, in Plots, the program offers us the most relevant plot, which is the object space (common space). In addition, we could request other plots that permit us to observe the model's degree of fit. As regards Output, the program provides us by default with the coordinates of the objects (in this case, cities) and the model's goodness-of-fit indices, as can be seen in Table 2.

COMPUTER PROGRAMS FOR MULTIDIMENSIONAL SCALING

There is a long list of computer programs for the solution of multidimensional scaling problems. We solved all the problems mentioned here using the PROXSCAL procedure, implemented in SPSS, but the same statistical package offers another procedure, called ALSCAL, which also permits the solution of multiple multidimensional scaling problems. In order to accede to this procedure the user must follow the sequence Analyze/Scales/Multidimensional scaling (ALSCAL). All the problems we have dealt with here could equally have been solved with ALSCAL.

In Table 5 we provide a short list of the computer programs currently on the market. In addition to those already mentioned, PROXSCAL and ALSCAL, multidimensional scaling problems can be solved with other programs, such as GGVIS, PERMAP, MULTISCALE or NewMDSX. GGVIS and PERMAP share the property of being interactive, and are available free of charge via Internet. MULTISCALE also has the advantage of free download from the Web, but it is difficult to use. Its author, Ramsay, is a highly prestigious figure in the history of multidimensional scaling. Finally, NewMDSX is, in reality, a program package that permits the solution of multidimensional scaling problems and of related types of problem.

If readers would like further information on computer programs and, more generally, on the history, models and methods of multidimensional scaling, as well as on its multiple application possibilities in Psychology and related matters, they might wish to consult the book by Borg and Groenen (2005) – the most recent manual on multidimensional scaling – or the works of Kruskal and Wish (1978), Arabie, Carroll and DeSarbo (1987), Green, Carmone and Smith (1989), or Arce (1993, 1994). For examples of applications, see Wish, Deutsch and Kaplan (1976) or Sabucedo and Arce (1990).

TABLE 5
COMPUTER PROGRAMS FOR MULTIDIMENSIONAL SCALING

NAME	AVAILABILITY	REFERENCES
PROXSCAL	In SPSS http://www.spss.com/	Commandeur & Heiser (1993), Meulman, Heiser, & SPSS (1999), De Leeuw & Heiser (1980)
ALSCAL	In SPSS http://www.spss.com/	Takane, Young, & De Leeuw (1977)
GGVIS	Free via Internet http://www.ggobi.org E-mail: ggobi-help@ggobi.org	Buja & Swayne (2002)
PERMAP	Free via Internet http://www.ucs.louisiana.edu/~rbh8900 E-mail: ron@heady.us	Ron B. Heady, University of Louisiana, Lafayette, USA
MULTISCALE	Free via Internet ftp://ego.psych.mcgill.ca/pub/ramsay/multiscl/ or from the author, Prof. James O. Ramsay, e-mail: ramsay@psych.mcgill.ca	Ramsay (1977)
NewMDSX	http://www.newmdsx.com/Coxon (2004)	Coxon (2004)

ACKNOWLEDGEMENTS

This research was carried out with the support of the General Directorate of Research, Development and Innovation of the Galician Regional Government (*Dirección Xeral de Investigación, Desenvolvemento e Innovación de la Xunta de Galicia*) (PGIDIT06PXIB211187PR).

REFERENCES

Arabie, P., Carroll, J.D., & DeSarbo, W.S. (1987). Three-way scaling and clustering. Newbury Park, CA: Sage.

Arce, C. (1993). *Escalamiento multidimensional: una técnica multivariante para el análisis de datos de proximidad y preferencia [Multidimensional scaling: a multivariate technique for the analysis of proximity and preference data].* Barcelona: Promociones y Publicaciones Universitarias (PPU).

Arce, C. (1994). *Técnicas de construcción de escalas psicológicas [Techniques for the construction of psychological scales].* Madrid: Editorial Síntesis.

Bennett, J.F., & Hays, W.L. (1960). Multidimensional unfolding: determining the dimensionality of ranked preference data. *Psychometrika*, 25, 27-43.

Borg, I., & Groenen, P.J.F. *Modern multidimensional scaling. Theory and applications.* New York: Springer.

Buja, A., & Swayne, D.F. (2002). Visualization methodology for multidimensional scaling. *Journal of Classification*, 19, 7-44.

Carroll, J.D. (1980). Models and methods for multidimensional analysis of preferential choice (or other dominance) data. In E.D. Lantermann, & H. Feger (Eds.), *Similarity and choice* (pp. 234-289). Vienna: Hans Huber.

Carroll, J.D., & Chang, J.J. (1970). Analysis of individual differences in multidimensional scaling via N-way generalization of Eckart-Young decomposition. *Psychometrika*, 35, 283-319.

Commandeur, J.J.F., & Heiser, W.J. (1993). Mathematical derivations in the proximity scaling (PROXSCAL) of symmetric data matrices. Tech. Rep. No. RR-93-03. Leiden, The Netherlands: Department of Data Theory, Leiden University.

Coxon. A.P.M. (2004). Multidimensional Scaling. In M.S. Lewis-Beck, A. Bryman, & T. F. Liao (Eds.), *The Sage Encyclopedia of Social Science Research Methods*. Thousand Oaks: Sage.

De Leeuw, J., & Heiser, W.J. (1980). Multidimensional scaling with restrictions on the configuration. In P.R. Krishnaiah (Ed.), *Multivariate analysis* (Vol. V, pp. 501-522). Amsterdam, The Netherlands: North-Holland.

Green, P.E., Carmone, F.J., & Smith, S.M. (1989). *Multidimensional scaling. Concepts and applications*. Boston: Allyn and Bacon.

Kruskal, J.B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29, 1-27.

Kruskal, J.B. (1964b). Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 29, 115-129.

Kruskal, J.B., & Wish, M. (1978). *Multidimensional scaling*. Newbury Park, CA: Sage.

Meulman, J.J., Heiser, W.J., & SPSS (1999). *SPSS Categories 10.0*. Chicago: SPSS.

Ramsay, J.O. (1977). Maximum-likelihood estimation in multidimensional scaling. *Psychometrika*, 42, 241-266.

Sabucedo, J.M., & Arce, C. (1990). Types of political participation: a multidimensional analysis. *European Journal of Political Research*, 20, 93-102.

Shepard, R.N. (1962). The analysis of proximities: multidimensional scaling with an unknown distances function (I and II). *Psychometrika*, 27, 125-139, 219-246.

Takane, Y., Young, F.W., & De Leeuw, J. (1977). Nonmetric individual differences multidimensional scaling: an alternating least-squares method with optimal scaling features. *Psychometrika*, 42, 7-67.

Torgerson, W.S. (1958). *Theory and methods of scaling*. New York: Wiley.

Tucker, L.R. (1960). Intra-individual and inter-individual multidimensionality. In H. Gulliksen, & S. Messick (Eds.), *Psychological scaling: theory and applications* (pp. 155-167). New York: Wiley.

Wish, M., Deutsch, M., & Kaplan, S.J. (1976). Perceived dimensions of interpersonal relations. *Journal of Personality and Social Psychology*, 33, 409-420.