

Using Trajectory Mapping to Analyze Musical Intervals

Stephen A. Gilbert

Dept. of Brain & Cognitive Sciences
MIT, E10-120
Cambridge, MA 02139
stephen@ai.mit.edu

Whitman Richards

Dept. of Brain & Cognitive Sciences
MIT, E10-120
Cambridge, MA 02139
whit@media.mit.edu

Abstract

Cognitive scientists have often pondered the question of perceptual spaces, that is, the question of how a certain gamut of familiar stimuli might be organized in the mind. We present Trajectory Mapping as an alternative clustering method to the traditional algorithm of Multi-Dimensional Scaling. We suggest that given data about the relationships among stimuli, Multi-Dimensional Scaling provides the one type of information (geometric), while Trajectory Mapping offers a second type (relational). As an illustration we present the initial results of applying both clustering techniques to subjects' perceptions of musical intervals. While an interpretation of the Multi-Dimensional Scaling requires *a priori* knowledge of music theory, Trajectory Mapping directly reveals the music theory that has been internalized by subjects.

Introduction

Cognitive scientists have often interested themselves in the question of perceptual spaces, that is, the question of how a certain gamut of familiar stimuli might be organized in the mind. Colors or textures, for example, can be organized with a variety of schema, and one might well wonder which of these the human being employs. Researchers in this field have often relied on standard clustering algorithms such as Multi-Dimensional Scaling to reveal the cognitive structures. This paper will point out a salient deficiency in such algorithms and offer an alternative approach called Trajectory Mapping (Richards & Koenderink, 1993). We will then illustrate the differences between the approaches by applying each to musical intervals.

To provide a conceptual comparison of a traditional clustering approach and Trajectory Mapping (TM), we offer an example involving 10 subway stations in the Boston subway system. If one rides a subway regularly, two types of information about the stations are acquired: the approximate distances between them (independent of their connections), and their placement along each of the various routes. The first type of information can be called **geometric**, while the second can be called **relational**. Figure 1, a map of the stations' locations and their connections, contains both types of information.

The MDS Technique

Multi-Dimensional Scaling (MDS), first discussed by Torgerson (1959), Shepard (1962), and Kruskal (1964), is a classic example of a clustering algorithm. When applied to our subway example, MDS derives the geometric

information mentioned above, namely the actual locations of the stations (up to a scaling factor). The input to MDS is a matrix of the distances between all the stimuli, and the output is the spatial arrangement of the stimuli that best fits the demands of the matrix. Note that MDS does not concern itself with the actual links between the stations.

Several noteworthy assumptions must hold in MDS. The distance matrix must be symmetric, that is, the distance from A to B is the same as from B to A. Also, it must be reasonable to assume that the stimuli could lie in a space in which a uniform distance metric holds (typically Euclidean). Furthermore, the researcher chooses the dimensionality of the fitting space. Although the "stress" parameter provides a quantitative assist in choosing the dimensionality, there is no qualitative indicator that derives it from the inherent structure in the stimuli.

The TM Technique

Trajectory Mapping focuses on relational information, such as the organization of our subway stations along various routes. While MDS is an analysis algorithm that can be applied to a distance or similarity matrix collected under a variety of experimental paradigms, TM describes not only the heuristic for data analysis, but also the experimental paradigm.

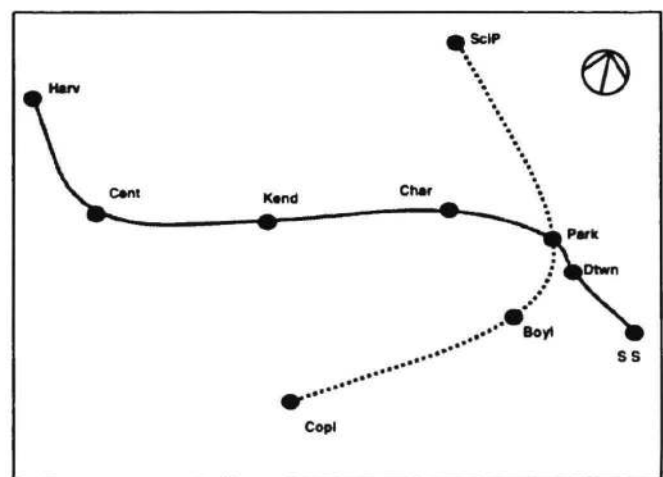


Figure 1: This crude map of a part of the Boston subway system contains two kinds of information. The geometric information comes from the physical distances between the subway stations. The relational information comes from the routes which link the stops together.

Figure 2: These TM data for the subway station example illustrate the Extrapolant—A—Interpolant—B—Extrapolant form of TM quintuplets.

ex	A	int	B	ex	ex	A	int	B	ex
—	S S		Copl	—	—	Harv	Park	S S	—
—	Cent		Boyl	—	—	Cent		SciP	—
Harv	Cent	*	Kend	Park	—	Char		Copl	—
Kend	Char	Park	Boyl	Copl	—	Kend		Boyl	—
—	SciP	Park	Copl	—	Harv	Kend	Park	Dtwn	S S
Harv	Cent	Kend	Char	Dtwn	Harv	Cent	Kend	Park	S S
Harv	Kend	Char	Park	S S	S S	Dtwn	Park	Boyl	Copl
Harv	Kend	Park	S S	—	Kend	Char	Park	Dtwn	S S
S S	Dtwn	Park	SciP	—	—	Kend		SciP	—
Cent	Kend	*	Char	Park	—	Harv		Boyl	—
—	Harv	Char	Dtwn	S S	—	S S		Boyl	—
—	Harv	*	Cent	Kend	—	SciP	Park	Boyl	Copl
Kend	Char	Park	SciP	—	Char	Park	*	Dtwn	S S
—	Harv	Kend	Char	S S	—	Harv		SciP	—
Kend	Char	*	Park	Dtwn	Harv	Cent	Char	S S	—
Harv	Cent	Char	Dtwn	S S	Park	Dtwn	*	S S	—
—	S S		SciP	—	Kend	Char	Park	S S	—
—	Copl	*	Boyl	Park	—	Cent		Copl	—
—	Harv		Copl	—	—	Dtwn		Copl	—
SciP	Park	*	Boyl	Copl	Boyl	Park	*	SciP	—
SciP	Park	Boyl	Copl	—	—	Harv	Cent	Kend	Park
—	Harv	Kend	Park	S S	Kend	Park	Dtwn	S S	—
—	Kend		Copl	—					

— dead end * feasible, no sample | infeasible

In the TM methodology a subject initially surveys the entire range of stimuli. If the experiment is visual, the subject might see a computer screen of images or a table full of picture cards. In each trial two stimuli are designated **A** and **B**. The subject is asked to note a feature that differs across the two and to choose a sample from the remaining stimuli which shows the result of extrapolating that feature further. In the case of colors, a subject given red and then pink might extrapolate to white. The subject also chooses an extrapolant going from **B** to **A** and an interpolant that would fit well between them. We ask the subject to use the same changing feature in each choice.

In the subway example, one might give a subject two station names and ask which station comes next in her mental subway map. Figure 2 contains possible subway data. Note that although the subway system is an object which one can describe objectively, we are not looking for a factual description with TM; we hope instead to derive each subject's perceptual map of the system.

There are three special cases that arise in this paradigm, as shown in the data. The first occurs when **A** and **B** are so dissimilar that the subject feels uncomfortable choosing a feature that smoothly varies between the two. We call this the "infeasible" case, and an example might be two subway stations that are at the far ends of two different subway lines (connected too indirectly). In the "feasible, no sample" case the subject can easily imagine an appropriate extrapolant or interpolant, but finds that the experimenter has not included it in the sample set. The "dead end" case occurs when either **A** or **B** represents an extreme in the variability range of the chosen feature. In the subway example, **A** or **B** would lie at the end of a line.

The analysis part of TM consists of outlining the dominant patterns in the data. One may divide the Extrapolant-A-Interpolant-B-Extrapolant quintuples into three overlapping triples, for example, and then identify the most prevalent triples across the entire data. Dominant triples that overlap could then be joined to illustrate the salient cognitive pathways through the space. In general, one seeks to highlight the common paths in the data. The exact approach taken can depend on the data set itself.

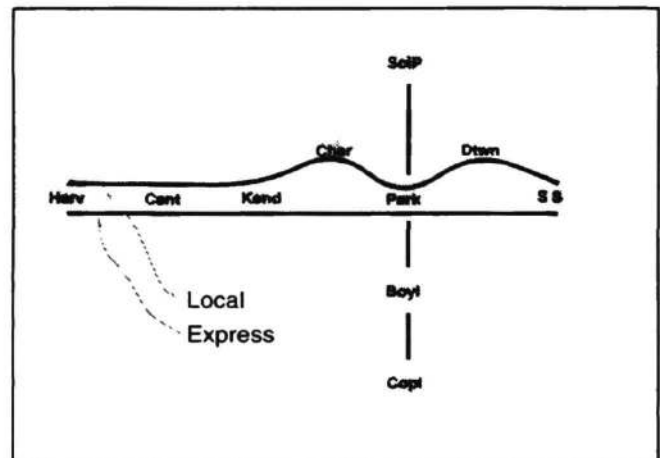
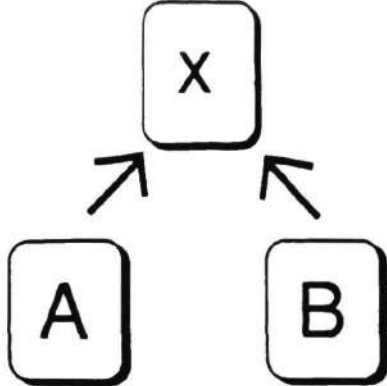


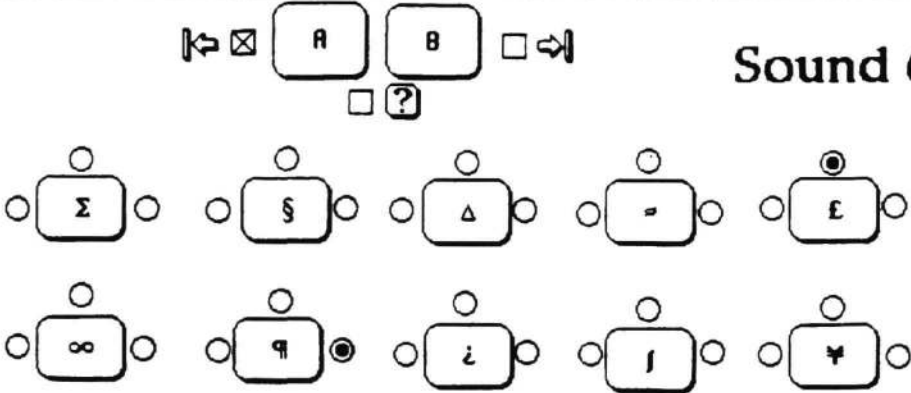
Figure 3: This Trajectory Map shows the relational nature of TM output. The "Local" and "Express" pathways represent different features, e.g. "all the stops on the red line" or "my favorite restaurant stops on the red line."

#2



Is X more similar to A or B?
Click the appropriate arrow.

Sound 6



? ?
 No Interpolant Available

Which 3rd notes are appropriate?



 Go On

Figure 4: The upper interface collects MDS data using the Method of Triads. The X, A, and B buttons each play an interval upon the subject's command. The subject must choose whether A or B is more "harmonically similar" to X by clicking the appropriate arrow. The lower interface collects TM data. The 12 buttons each play an interval. The subject thinks of a feature that varies across the two primary intervals, A and B, and extrapolates by picking a third interval that would best continue that change (¶). The subject also chooses an interpolant in the other direction (none possible here) and an interpolant (£).

In the case of subway stations, doing TM with an experienced commuter would likely yield a trajectory map very similar to the schematic subway maps seen inside the stations themselves, diagrams that show the routes among the stations with minimal indication of their true geography. Figure 3 shows a trajectory map of the example data. Note that the trajectory map has two pathways that travel between *Harv* and *S S*. Since each trajectory represents a subject's moving among stimuli using a certain feature, different trajectories that thread similar stimuli can indicate different contexts in which the subject might move among stimuli, i.e. different features of the stimuli. In the subway example, the "Local" could represent a subject's pedantic recalling of the subway map, while the "Express" might represent the subject's individual traveling map; he rarely stops at Char or Dtown stations.

Organizing Musical Intervals

Now that we have briefly introduced the clustering algorithms, we offer their different approaches to discovering subjects' cognitive organization of musical intervals. Various studies have examined the relationship between the perception of notes and their tonal context or how a perceiver organizes notes within a single tonal context (e.g. Krumhansl & Shepard, 1979; Dowling 1986; and Krumhansl 1979, 1990). Like some of these studies, we ask what features or organizing principles a subject uses to group musical intervals, and specifically, we ask what TM and MDS can tell us about these features.

We define a musical interval as a pair of different notes from the traditional Western 12-tone octave in which the note of lower pitch occurs first. This definition gives us a total of 12 stimuli, and we denote them as m2, M2, m3, M3, P4, T, P5, m6, M6, m7, M7, and O. The "m" means "minor," the "M" means "major," the "P" means "perfect." "M2," for example, represents a rising interval of a major second. The "T" is the tritone (the physical middle of the octave), and the "O" is the octave. Within a given trial, we begin all stimulus intervals on the same note to maintain a consistent tonal context within each trial. Across trials, we vary the beginning note within an octave range.

The MDS Experiment

We exposed a musical and non-musical subject to 660 trials using a Yamaha DX-7 synthesizer and a Macintosh running Hypercard and MIDI. The upper half of Figure 4 presents the interface. We presented the subjects with three buttons A, B, and X, each of which played an interval when clicked with the mouse. We asked subjects to indicate whether X was more "harmonically similar" to A or B by clicking the appropriate arrow. Subjects could listen to the intervals as many times as they wanted by clicking the buttons. We presented each of the 12 stimuli behind the X matched with all the possible pairs of remaining stimuli (660 trials).

To calculate the similarity matrix from the triad data, we followed Torgerson's instructions for "The Complete Method of Triads" (1952); the number of times that each pair was chosen became its similarity measure. We then computed the MDS space with KYST2 (Bell Labs) using

stress formula two and the default options. Figure 5 holds the average of the two dimensional results from a musical and non-musical subject (stress = 0.4).

The output of the MDS illustrates the pros and cons of the approach. It appears that the axes of the plot may represent musically significant variables, such as interval size (growing with the y-axis) and traditional tonal dissonance (decreasing with the x-axis), but there is nothing explicit in the space or in the MDS algorithm that points to such axes. Further work, such as experimental validation of hypothesized axis extrema, must be done to support any particular selection of axes. Thus, the interpretation of an MDS plot requires external *a priori* knowledge about the stimuli.

We included the corresponding MDS results from Krumhansl's (1979) pitch judgments given a tonal context set by a scale or chord (stress = 0.2). It is perhaps worthwhile to note how similar the results are, and that the differences might be explained by the different paradigms. One might imagine that while an M7 would seem consonant after hearing a scale (Krumhansl's data), hearing it as an interval among other intervals with the same tonic might make it seem dissonant (our data). The reverse applies to the m7 and m3 since their upper notes are not in the working key, while they do play a role as intervals in familiar chords of the working key.

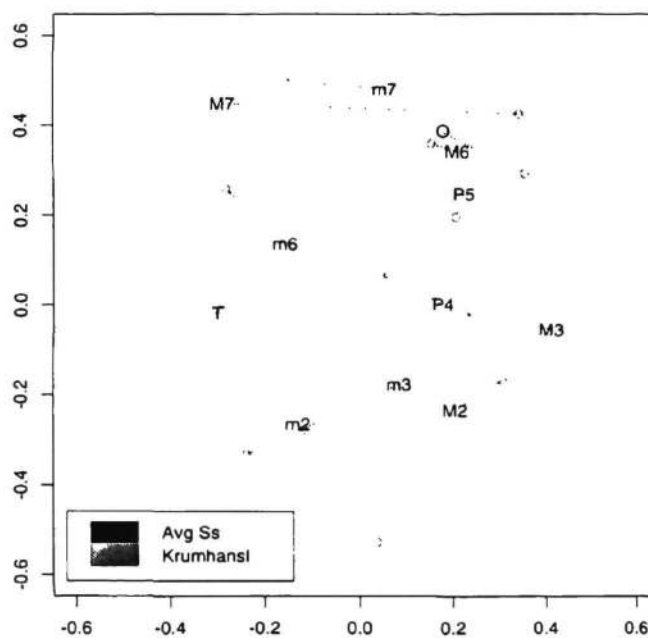


Figure 5: The above graph shows the average of the MDS outputs from the data of a musical and a non-musical subject (in black). Note that although the axes may appear to represent musically significant features, such as interval size (growing with the y-axis) and traditional tonal consonance (growing with the x-axis), one must experiment further to verify this interpretation. The corresponding MDS results from Krumhansl (1979) are shown in gray to point out the gross similarity of results despite different methods of setting tonal context.

ex	A	int	B	ex	ex	A	int	B	ex
—	m2	*	M2	m3	—	P4		M7	—
—	m2		m7	—	—	P4	M6	O	—
—	m2		M7	—	P4	T	*	P5	m6
m2	M2	M3	P4	P5	—	T		M7	—
—	M2		m6	—	P4	P5	*	m6	m7
—	M2		m7	—	P5	m6	*	M6	m7
—	M2		M7	—	P5	m6	m7	M7	O
M2	m3	*	M3	P4	P5	M6	*	m7	O
—	m3		M7	—	M6	m7	*	M7	O
—	M3		M7	—	m6	m7	*	M7	O
—	P4	P5	M6	O	m7	M7	*	O	—
—	P4		m7	—					

— dead end * feasible, no sample | infeasible

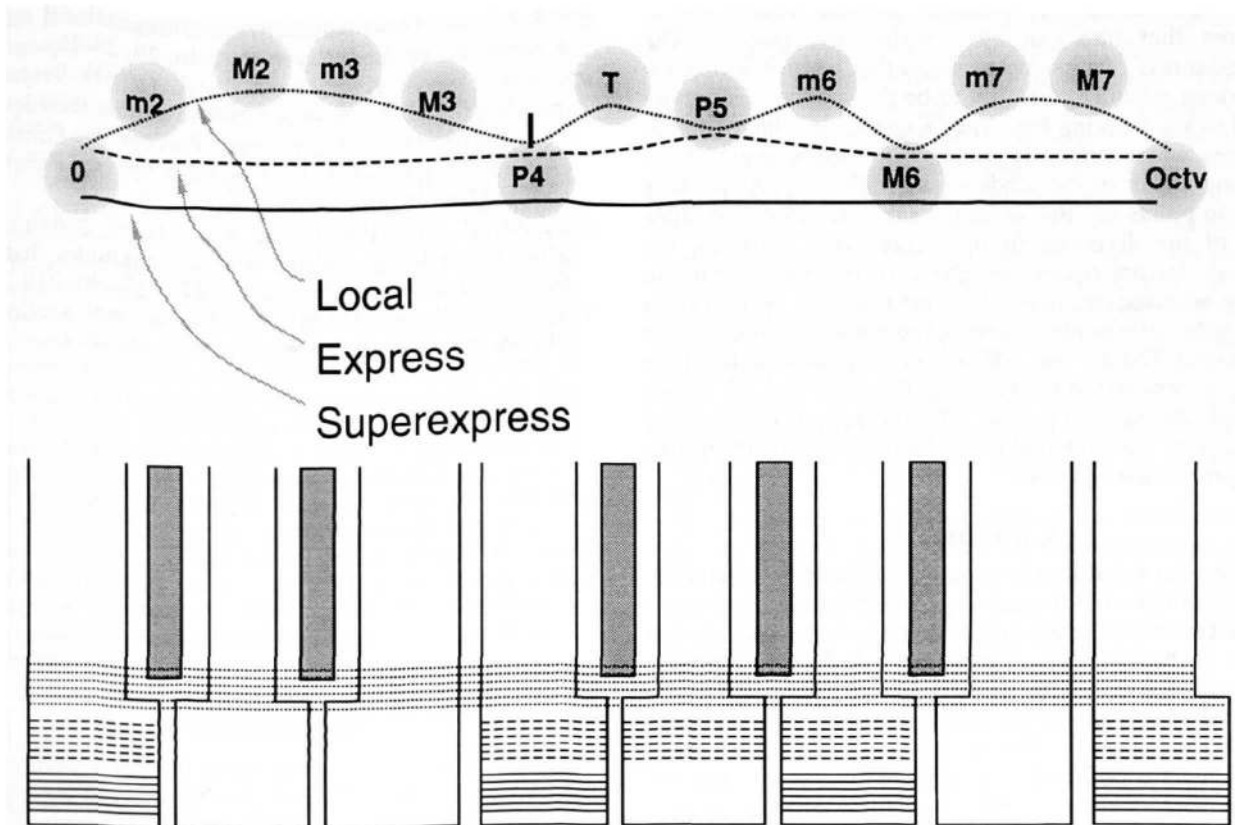


Figure 6: These TM data are the common quintuplets from a musical and a non-musical subject. The resulting map illustrates TM's ability to reveal features based on data alone; the data reveal three main pathways in the subjects' perceptual spaces. In some sense the pathways confirm traditional music theory, e.g. the "Superexpress" represents the subdominant major chord of the working key. Note also that subjects do not weight the 12 equal divisions of the octave linearly.

The TM Experiment

We exposed the same musical and non-musical subjects to 66 trials using similar hardware and software as in the MDS experiment. The Hypercard interface (shown below in Figure 4) offered the subject two primary buttons, A and B, that played the two intervals of interest. A subject was asked to think of a feature that changed between them and then to extrapolate by picking a third interval from the remaining 10 that would best continue that change. We also asked for an extrapolant in the other direction and an interpolant, with the requirement that all five intervals in the resulting quintuple formed a coherent sequence in both directions, meaning that the same feature varied throughout.

To show the capability of TM, we chose to analyze those quintuples of data which were identical across subjects. These data appear in Figure 6. Quintuples were linked by joining at least two or more overlapping members of each quintuple. These linked quintuples determine the paths along the octave. The trajectory map in Figure 6 was then derived using only the paths that included the most frequent links, also taking into account the special case markers (see Richards & Koenderink for detailed procedure).

Just as in the subway example, the data reveals several features that one can vary across the stimuli. The "Superexpress", for example (named thus because it contains the fewest stimuli), turns out to be the subdominant major chord of the working key. The "Express" is a larger subset of notes which primarily suggests the major scale of the working key or of the subdominant major key. As a whole the map points out that although subjects can differentiate each of the divisions of the octave (the "Local"), the intervals are not equally weighted. (If one knew no music theory whatsoever, one might assume that an octave, a physically defined phenomenon, has equal and independent divisions.) The P4 and M6 are pivot points on all three paths, as opposed to the centered Tritone or the 1/3 octave markers, the M3 and the m6. TM thus reveals that some of the subjects' own internal music theory corresponds in part with general music theory.

Conclusion

Having seen the different types of information offered by MDS and TM, it is natural to point out that both methods can be combined if one's goal is finding a representation that offers both geometric relationships and the contextual connections (as does the original subway map in Figure 1). What we would like to emphasize, however, is that TM can offer an alternate method to those researchers who have traditionally used MDS to establish a geometric space and then applied various other techniques of estimating what the primary relational axes of the space are (the main features or contexts). If one's main interest is these features, TM offers the efficient solution of eliciting them directly from the data. Richards & Koenderink have used TM in studies with color and visual textures (1993), and we'd like to conclude that it could assist music studies as well; TM not only reveals the cognitive relationships among stimuli without requiring the symmetry and uniform distance metric assumed for

traditional MDS, but also can help discern which aspects of traditional music theory have been absorbed by the minds of musicians and non-musicians.

Acknowledgment

Supported by NSF grant IRI 89-00267 to Whitman Richards and NIH Training grant T32-GM07484 to Stephen A. Gilbert. We thank Kathryn Vaughn for helpful discussion of these ideas.

References

- Dowling, W. J. (1986) Context effects on melody recognition: Scale-step versus interval representations. *Music Perception*, 3, 281-296.
- Krumhansl, C. L. (1979) The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11, 346-374.
- Krumhansl, C. L. (1990) *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.
- Krumhansl, C. L. & Shepard, R. N. (1979) Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception & Performance*, 5, 579-594.
- Kruskal, J. B. (1964) Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29, 28-42.
- Richards, W. & Koenderink, J. J. (1993) Trajectory mapping ("TM"): A new non-metric scaling technique. in *Proc. European Conf. on Visual Perception, Edinburgh* (Abstract). *MIT Center for Cognitive Science Paper #48 and AI Memo #1468*, 1994.
- Shepard, R. N. (1962) The analysis of proximities: Multidimensional scaling with an unknown distance function. I. & II. *Psychometrika*, 27, 125-140, 219-246.
- Torgerson, W. S. (1952) Multidimensional scaling: I. Theory and method. *Psychometrika*, 4, 401-419.