Obtaining Psychologically Motivated Spaces with MDS

Mikael Johannesson

Lund University Cognitive Science
Kungshuset, Lundagård
S-222 22 LUND, Sweden
and
Department of Computer Science
University College of Skövde, Box 408
S-541 28 SKÖVDE, Sweden

mikael.johannesson@ida.his.se

Abstract: The main purpose with this paper is to describe how a psychologically motivated conceptual space can be obtained with MDS (multidimensional scaling) and how it can be expressed in terms of a more primitive "physical" (or mathematical) one. The idea is demonstrated practically with the aid of two experimental pilot studies. The paper is concluded by a critical discussion of the method used.

INTRODUCTION

Always when modelling cognition, the representation is of significant importance. It is the underlying representation medium, together with the operations that work upon it, that constrains what can be modelled. Throughout the history of cognitive science, a tremendous amount of different types of representations have been used and these could be categorized in several ways depending on what properties is focused upon. For the purpose of this paper, a relevant distinction is between *linguistic* (symbolic) and *non-linguistic* based representations. Linguistic representations, which historically have been most common, are required to be expressible in linguistic code, whereas non-linguistic are not. For a general discussion of representational issues (Palmer, 1978) is strongly recommended.

The type of representation of interest here is that of *conceptual spaces* (Gärdenfors, 1992), which is subsumed by the more general framework often referred to as *multidimensional spaces*. Conceptual spaces are limited to representing concepts, hence the term *conceptual*.

A conceptual space consists of a number of *quality dimensions*, each representing some quality of a concept. A basic metaphor is that mental objects could be represented as *points* or *regions* in a conceptual space and that similarity is reflected by distance relationships between them. The main purpose of this paper is to, with the help of two experimental studies, present how a conceptual space expressed in terms of a "physical" (or mathematical) one could be obtained by using multidimensional scaling.

Quality dimensions are taken to be *cognitive* and *infralinguistic*, i.e. the qualities they represent are not required to be expressible in linguistic code. Various quality dimensions may differ in a number of ways, which are probably not completely exclusive:

- *nature of the scale*: The underlying scale may be of nominal, ordinal, interval or quote nature. A nominal scale merely divides objects into classes (e.g. sex), whereas the other scales are (at least) ordered (e.g. pitch).
- *metrical structure*: Some quality dimensions are discrete to their nature, others are continuous.

Lund University Cognitive Studies – LUCS 45. ISSN 1101–8453.

- *range of the scale*: A dimension could, for example, be isomorphic to the half-line of non-negative numbers (e.g. weight), or be isomorphic to the line of real numbers (e.g. time).
- *sensory-nonsensory*: A specific quality dimension could be anywhere within the range from a close relation with what is produced by our sensory receptors (e.g. pitch) to be of an abstract and non-sensory character (e.g. dangerous).
- *origin*: Some dimensions seem to be innate and to some extent hardwired in our nervous systems (e.g. pitch). Others are probably learned. Some dimensions are culturally dependent, e.g. some cultures conceive time as linear, but some conceive it as circular.

Sometimes, a distinction could be made between *scientific* and *psychological* interpretations of quality dimensions. One example is the difference between the ideal euclidian 3-D space and our psychological visual space. In our psychological visual space, the vertical dimension is treated differently from the horizontal dimensions, much because of gravity. In the ideal 3-D euclidian space, all three dimensions are treated in the same way. Psychological dimensions are discussed in some detail in e.g. (Gärdenfors, 1992). In this paper, the operational definition that *psychological quality dimensions are the dimensions that could be used for describing peoples ratings of similarity/dissimilarity between stimuli*, is adopted. All psychological quality dimensions used, together constitutes the psychological conceptual space for the stimuli. The terms “psychological conceptual space”, “psychological space” and “psychologically motivated space” are used synonymously.

As already mentioned, the assumption underlying multidimensional space-like representations is, that if mental objects are located at unique locations in a space (e.g. points), the psychological similarity is reflected by distance relationships between them. This means, that the closer two objects are (i.e. the shorter the distance), the more similar they are. So, if the underlying psychological quality dimensions are known, then the psychological space is known. Furthermore, if the locations of mental objects (in this case, stimuli) also are known, then, if the *distance function* is known, the *similarity between objects could be calculated*. Of course, there is a tight relationship between the underlying scales and what could be inferred from distances.

Once there is a way of calculating similarities between locations in a conceptual space, more complex things could be represented. For example, there are several different models that operate upon multidimensional spaces for predicting concept formation behaviour. All of them have in common the assumption that categorization is based upon similarity between concepts and that they are exemplar based. Some examples of models are *Nearest neighbour* (NN), *Average Distance* (AD), *Prototype Voronoi Tessellation* (PV) and *Generalized Voronoi Tessellation* (GV).

According to NN, a new object belongs to the same category as its closest neighbour does. According to AD, a new object belongs to the same category as its closest hypothetical average member does. Both PV and GV resemble NN to some extent. According to PV, a new object belongs to the same category as its closest *prototypical point*, or object, does. A prototypical point could be taken to be the most central member of the category. According to GV, a new point belongs to the same category as its closest *prototypical region* does. Everything within the prototypical region could be the most central member to the category together with a radius that corresponds to the variation of what is considered to be typical.

No matter which model is used, if we want to evaluate its describing power, the underlying space and its dimensions are critical. The question is, how do we get knowledge about this space?

Abstract dimensions could very well fit the operational definition of psychological quality dimensions adopted here. However, a space constituted by such dimensions could be impractical for evaluating concept formation models that work upon it, the reason being that the dimensions may be hard to manipulate. For this reason, the aim of this paper is to present how spaces that are as simple as possible w.r.t. what dimensions are used, rather than the number of dimensions used, could be obtained.

The psychological quality dimensions for describing concepts or stimuli, and knowledge of these, could be obtained in several ways. In psychophysics, where the general aim is to quantify the relationship between physical characteristics and psychological sensations of stimuli, one normally studies one dimension at a time in order to find interesting magnitudes like absolute thresholds, difference tresholds and points of subjective equality. A basic assumption is that the physical and the psychological dimensions are the same. When the task is more molar (opposite to molecular), e.g. to study more complex behaviour like categorization, recognition etc., other analyses are required. For example, unidimensional analyses are not sufficient when more complex stimuli are used. Two examples of simple methods for obtaining psychological quality dimensions is to assume that it correspond directly to the multidimensional “physical” space, or to ask the test subjects which dimensions they use, for example, when judging the similarity within a set of stimuli. These methods are not purely advantageous, however, one reason being that the psychological space may not have a direct relation to the “physical” one. Also, subjects may not know which “physical” dimensions (mathematical dimensions which describe some physical characteristic of an entity) they really base their judgements on or how they relate to similarity judgements. There are methods which do not suffer from these specific problems, i.e. that allow the experimenter to find dimensions that could be used to constitute an underlying psychological space, without depending directly on the "physical" dimensions. Commonly, the dimensions searched for are often of a rather abstract nature, e.g. dimensions like "happy/unhappy", "social awareness" etc. Geometrical procedures for finding dimensions in psychological spaces or configurations are often called multidimensional scaling procedures, and will be the major method discussed in this paper.

MULTIDIMENSIONAL SCALING

Multidimensional scaling (MDS) is the universal name for a set of mathematical procedures which let the user represent stimuli in a spatial manner. The procedures are essentially concerned with finding a configuration of mental objects, concepts, stimuli etc., based on the perceived distances (*proximity data*) between them. An analogy is to create a table of distances between pairs of cities from a map. But instead of providing a table of distances, MDS-procedures does the opposite – they provide a map based on such a table. However, especially if the euclidian metric is used, neither the orientation of the resulting map nor the location of a specific city is known, the reason being that there is a 1:N mapping between a given set of distances and the corresponding spatial configurations.

Somewhat simplified, MDS-procedures work as follows: once a starting configuration based upon the dissimilarities/similarities (*proximity data*) is calculated, an iterative phase begins. During each iteration the distances between the points (stimuli) in the configuration are calculated and compared with the proximity data. If the differences are too big according to some criterion, the configuration is modified. The modification is based on the method of *steepest-descent*, which means that for a particular pair of points, the modification is proportional to the difference. When the configuration is sufficiently "good", the procedure terminates.

Metric and Non-Metric MDS

There are two main types of scaling procedures: *metric scaling* and *non-metric scaling*. Metric scaling assumes that the *numerical values* of the dissimilarities (or similarities) are of *significance*. Metric scaling procedures works upon some instance of the general Minkowski metric, which is defined by

$$d_{ij} = \left\{ \sum_{a=1}^{p} |x_{ia} - x_{ja}|^R \right\}^{1/R} \quad ; \quad 1 \le R \le N$$

where d_{ij} is the distance between points j and s. The far most common metric used is the *euclidian metric*, which is defined by Minkowski R = 2. Another well known metric is defined by Minkowski R = 1, and is called the "*city-block*" or "*Manhattan*" metric. An important difference between the two metrics is that a configuration based on the euclidian metric is rotatable, whereas a configuration based on city-block, is not (in general, due to the fact that rotation – other than multiples of 90 degrees – does not preserve the original distances). Metric scaling is used when the level of measurement of data is interval or ratio. Comparing the configuration and proximity data is straightforward since the distances in the configuration could be directly compared to the proximity data.

In non-metric scaling, the exact numerical values of the distances are assumed to have little intrinsic meaning, it is rather the *relative order* which is considered to be important, and the level of measurement is assumed to be ordinal. For non-metric scaling, some monotone transformation has to be applied to the proximity data (which is often referred to as *disparities*) in order to allow the arithmetic operations necessary for comparing them with the configuration. In fact, non-metric MDS with monotonicity and a minimum of dimensions is generally enough to recover metric information from nonmetric information (see e.g. (Shepard, 1962)).

In short, the major differences between the two types of scaling are:

- Metric scaling tries to fit the dissimilarities/similarities to the distances in the configuration.
- Non-metric scaling tries to fit the rank order of the dissimilarities/similarities to the distances in the configuration.

Comparison Between Configuration and Empirical Data

The criterion for termination of the iterative process is usually based on some measure of goodness or badness of fit. The particular measure may differ between different programs. An example of a common such measure for ordinal proximity data is *stress*, which is a measure of departure from monotonicity. Of the several versions of stress, one is defined by

$$Stress = \left[\sum_{i}^{n} \sum_{j}^{n} (d_{ij} - \hat{d}_{ij})^2 \Big/ \sum_{i}^{n} \sum_{j}^{n} (d_{ij}^2) \right]^{1/2}$$

where d_{ij} is the distance between points i and j, and $\hat{d}_{ij}$ is the *disparity* (transformed dissimilarity) for stimuli i and j. Stress is a measure of badness of fit and is thus inversely related to goodness of fit.

Dimensionality of the Configuration

One of the main points with MDS is to find the *relevant* number of dimensions. There does not seem to be a "best way" for determining which number is relevant, but there are some heuristic principles (see e.g. (Kruskal & Wish, 1978), (Shepard, 1972)). A common way is to compare the goodness of fit for configurations of different dimensionality. There is no point comparing these measures directly, since the goodness of fit will be better with increasing dimensionality. Rather, the *changes* in goodness of fit between configurations is compared. If a change between two dimensions, say N and N-1, is substantial compared to the other changes, then N is usually considered as the "correct" number of dimensions.

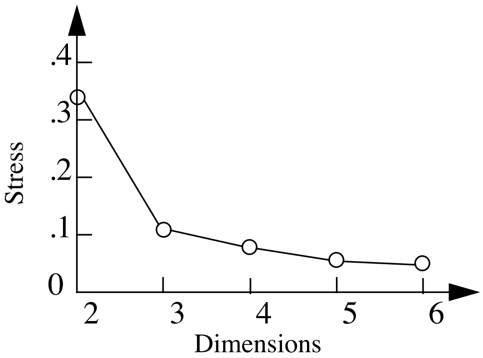


Figure 1. Changes in goodness of fit (Stress) with dimensionality.

In the constructed example in Figure 1. it is easy to see that the correct number of dimensions is three. Another way of expressing this is to say that the number of dimensions at the "*elbow*" is the correct one. In other words, the relevant number of dimensions, is the number of dimensions that provides the best change in goodness of fit compared to the next lower number. The guideline appears to be very simple, but there are some factors that make it a bit more complicated than is outlined here (see e.g. (Kruskal & Wish, 1978)). There also are heuristic principles which are not directly based on goodness of fit (see e.g. (Schiffman, Reynolds & Young, 1981)).

It is important to note that guidelines, as described in this paper and others, are merely guidelines. They have evolved from people's experience with MDS, and have been useful. They are, however, not guaranteed to find the number of dimensions actually used by subjects. No matter which guidelines are used, since MDS is used for understanding proximity data, one has to consider factors as interpretability, ease of use and stability. All these factors should affect the choice of dimensionality.

Interpretation of the Configuration

One of the most important and obvious methods for examining the configuration is to look at the arrangement of stimuli. What is searched for are lines that arrange stimuli which are maximally different from each other, such that these appear at polar ends of the line. The method, however, is limited to configurations of small dimensionality. Already with three dimensions the method can be quite difficult, depending on visual limitations, in four or more dimensions visualisation becomes impossible. Another problem is that it is difficult to say whether or not a dimension is genuine if the described relationship is not very strong. Especially in such cases, it is better to rely upon statistical techniques like (multiple) *linear regression*, which could be used to find the linear function over the co-ordinates (the *independent variables*) that best predict some characteristic of stimuli (the *dependent variable*). The *determination coefficient* R^2 (or multiple regression coefficient) provides a direct measure of how well the dependent variable can be predicted. According to (Kruskal & Wish, 1978), it is desirable that the determination coefficient is in the .90's for a good interpretation, but in some cases R^2 in the .80's or upper .70's has to suffice. Furthermore, a minimal requirement is a significance level at the .01 level or better.

The quote from (Shepard, 1972, p. 10): "*It is only in the cases of non-Euclidian metrics and other certain cases – ... – that the axes that come out of multidimensional scaling can be expected to be interpretable without further rotation*", gives a hint about when linear regression analysis is applicable. As a matter of fact, there is no use performing such an analysis on a configuration based on the city-block metric. The reason is that the space cannot be rotated to the eventual optimal axes. However, it is, of course, possible to examine the correlations between the outcoming axes and the dependent variables.

Even if the dependent variables are varied in several more or less different ways, e.g. ratings of stimuli on several different linguistically coded scales, there is no guarantee that the "proper" dimensions are in the set. This is because the dimensions must be known before the regression or correlation analyses. The problem can be reduced by using a range of possibly relevant dependent variables, analyse them, and calibrate the interpretation afterwards.

EXPERIMENTAL STUDIES

In order to demonstrate that MDS can be a useful tool for finding a mapping between a "physical" or mathematical description of a stimuli space and the corresponding psychological space, two experimental studies will be accounted for. The goal with each of them have been to find the underlying psychological space and describe it in terms of the “physical” parameters generating the stimuli. The more general goal is to demonstrate how psychologically motivated spaces could be obtained, spaces that distance based categorization rules etc. could work upon.

For the analysis of both the studies, the MDS-program KYST-2a (Kruskal, Young & Seery, 1977), have been used. Scaling analyses has been done on single matrices representing average judgements from subjects. Configurations based on both the euclidian metric and the city-block metric, have been analysed. In the following, KYST-2a will be referred to as KYST.

THE MOLLUSC SHELL STUDY

The study is similar to the pilot studies performed by Gärdenfors and Holmqvist (Gärdenfors & Holmqvist, 1994). The goal with the present study is the same as one of theirs: to find a psycho-physical space for mollusc shells. However, the studies differ in the methodology adopted.

Gärdenfors and Holmqvist collected similarity judgements from thirteen subjects by letting them rate how similar the middle shell (of three shells) was to the left and the right shells on a continous scale. The shell space used consisted of three dimensions V’, E’ and R (see below). In order to obtain a satisfactory description of subjects’ perceptions of shell forms, transformations in form of instances of Stevens power law, $d'(x) = k(d(x))^b$, were tested. The vectors (kV’,kE’,kR,bV’,bE’,bR) were varied so as to maximise the Pearson correlation coefficient between averaged test sheet distances (based on the similarity ratings) and the shell space distances.

Even though there are some resemblance between Gärdenfors and Holmqvist’s study and the present, they are hard to compare. This is mainly due to the fact that they aimed at deriving the constants of Stevens power law when the space consisted of three dimensions decided in advance.

Stimuli

The same stimuli as Gärdenfors and Holmqvist, i.e. computer generated pictures of mollusc shells, were used. An advantage of using shell shapes as stimuli is that people normally have quite little experience with them, but that they yet are rather “ecologically valid”.

The shells varied in three mathematical (or “physical”) dimensions (see Figure 2.): The rate E of whorl expansion, which determines the curvature of the shell, the rate V of vertical translation along the coiling axis and the expansion rate R of the generative curve of the shell.

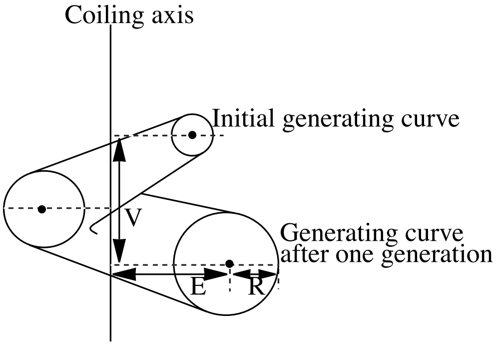


Figure 2. The dimensions generating a mollusc shell shape.

Gärdenfors and Holmqvist made no direct use of the “physical” parameters E and V. Rather, they used the transformations V’=V+R-1 and E’=E+R with the motivation that people do not look at the centre of the generating circle when estimating the vertical and horizontal growth rates of shells, but rather on the height and width of them.

Figure 3. presents the “physical” parameter space used in the present study. Figure 4. and Figure 5. show the resulting stimuli.

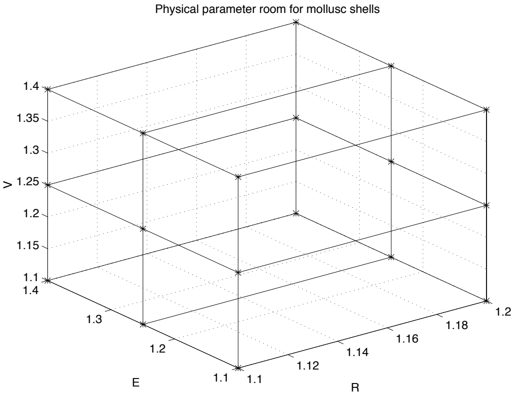


Figure 3. The underlying “physical” space for the mollusc shells.

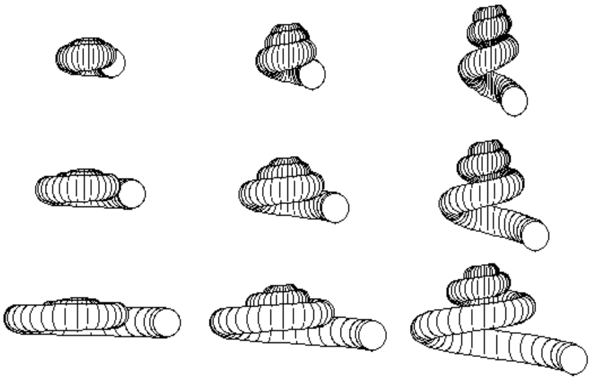


Figure 4. Mollusc shells with R = 1.10. Parameter V varies with the horizontal axis of the figure (1.10,1.25,1.40) and parameter E varies with the vertical axis (1.40,1.25,1.10).

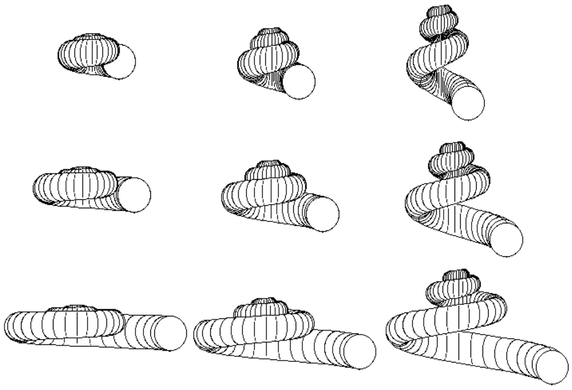


Figure 5. Mollusc shells with R = 1.20. Parameter V and E has the same variation as in Figure 4.

Method

The direct scaling data was collected during a session with PsyScope (Cohen, MacWhinney, Flatt & Provost, 1993), a program especially developed for presenting experiments on a computer. The session, programmed by Måns Holgersson at Lund University Cognitive Science and the author, mainly consisted of three different phases: an instruction phase, a presentation phase and a similarity rating phase.

Before a session the experimenter informed the subject personally that a typical session would take about 40 minutes[1], that there would be breaks[2] and that instructions would be given on the screen.

The Instruction Phase

The first screen contained general information about the test and instructions for the similarity rating part, all in Swedish. Subjects were told that the purpose of the study was to investigate how human subjects make similarity judgements between mollusc shells. They were also told to base their ratings on whatever attributes they felt was relevant, but as consistently as possible. The similarity ratings should be mapped to a nine-grade similarity scale, where “1” corresponded to “large similarity” whereas “9” corresponded to “large dissimilarity”. Subjects were instructed to not spend too much time on particular pairs, but rather try to make holistic judgements.

The Presentation Phase

First, subjects had each shell in the forthcoming session presented to them. During this presentation, shells were presented pairwise in a randomised order that was the same for all subjects. Each pair was showed for 5 seconds. The reason for presenting the stimuli before collecting data was that subjects should know the variation between stimuli, i.e. subjects were calibrated with respect to the variation of stimuli. Similarity ratings where the variation of stimuli is previously unknown to subjects, is probably likely to be more unstable until all stimuli have been presented.

The Rating Phase

This phase consisted of rating 153 different pairs[3] of mollusc shells with respect to the similarity between the members of a pair. Each screen (see Figure 6.) consisted of two shells in left - right order and a scale symbol bar marked with each digit from 1 to 9. It was also indicated that lower numbers corresponded to large similarity, whereas higher numbers corresponded to larger dissimilarity.

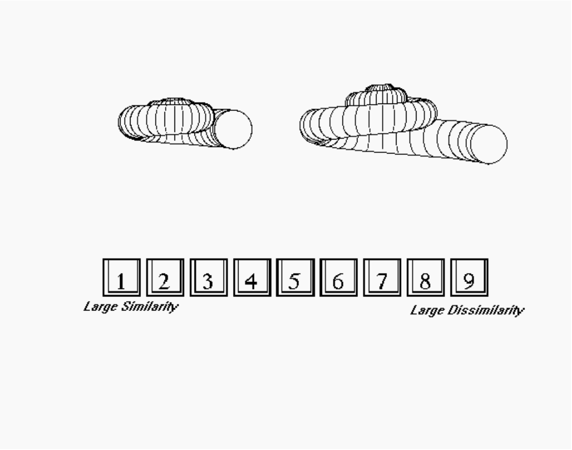


Figure 6. Example of a screen during the similarity rating phase.

In order to avoid systematic bias, the presentation of the pairs to judge was block-randomised with respect to the order of the pairs between subjects. Further, the presentation of a single shell was balanced in left-right-order, i.e. of the total 17 times each shell appeared, it appeared 9 times on one side and 8 times on the other.

Subjects were divided into two groups, group A and group B. Which subject that belonged to which group depended on the order in which the studies was performed. The first subject belonged to group A, the next to group B, and so on. The difference between the groups was in which left-right order a single pair was presented. All pairs that members of group A encountered as Stimuli x to the left and Stimuli y to the right, was presented as Stimuli y to the left and Stimuli x to the right for group B. The reason for using two groups was to be able to investigate if asymmetry in left-right order could be traced. However, such an analysis is beyond the scope of this paper.

---

[1] The collection of data, as described in this paper, actually took shorter time. The real session also consisted of an adjective rating part. This part was the last segment of the study, and could therefore have had no affect on the study as described here.

[2] The exact number of breaks throughout the whole session was 6. The breaks had no time limit.

[3] The number of unique pairs of N stimuli = $\frac{N \cdot (N-1)}{2}$

Subjects

Eleven subjects participated in the study without any credit. Six subjects were assigned to group A and five subjects to group B.

Results and Discussion

Number of dimensions and the Nature of the Scale

In order to find the “correct” number of dimensions for the mollusc shell space, stress versus dimensions was plotted for the six to one dimensional solutions (Figure 7.).

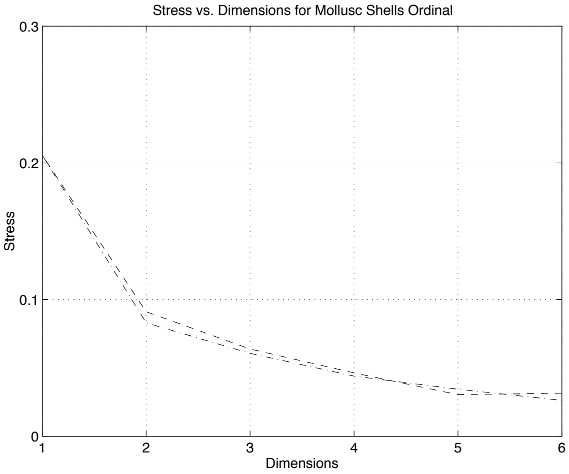


Figure 7. Stress vs. dimensions for mollusc shells using KYST euclidian metric (dash dotted) and KYST city-block metric (dashed).

In all solutions, stress increases smoothly with decreasing dimensionality from six to two dimensions, where there is an apparent elbow, suggesting a two dimensional solution.

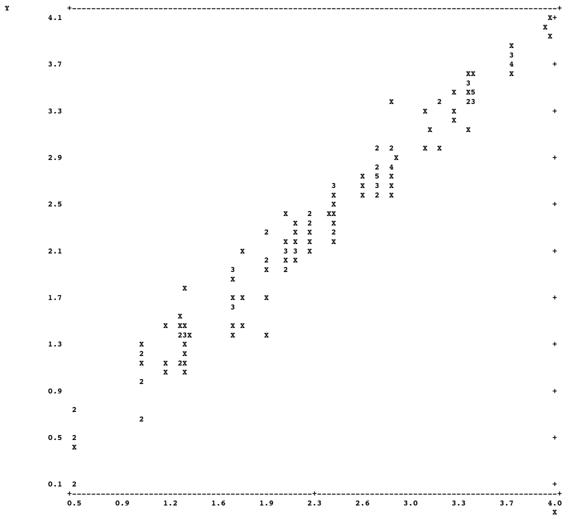


Figure 8. Scatter diagram provided by KYST for mollusc shells: distances (vertical) vs disparities (horizontal)

However, before starting the procedure of interpreting the configurations suggested, it is wise to examine the relation between the distances in stimulus space and disparities (transformed data). The scatter diagram in Figure 8. shows a plot of distances versus disparities. It is clear that the relationship is approximately linear. This means that the monotone transformation was linear and that metric scaling could be used instead. In light of this finding, the scalings were repeated with metric scaling instead of non-metric scaling[4].

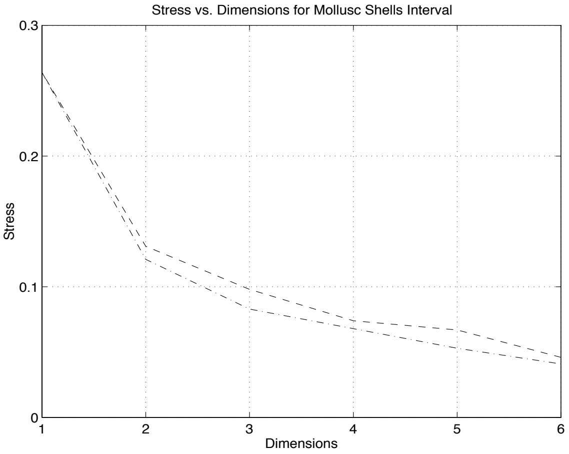


Figure 9. Metric scaling (interval data): Stress vs. dimensions for mollusc shells using KYST euclidian metric (dash dotted) and KYST city-block metric (dashed).

The stress vs. dimensions plot in Figure 9. is somewhat different compared to the one in Figure 7.: the stress values are generally higher. This could not be viewed as a disadvantage for metric scaling in this case since stress values always are higher for the interval level compared to the ordinal level for same data. There still seem to be an elbow at two dimensions, even though it is less clear than above. Therefore, the choice between a two- and a three-dimensional configuration could be better grounded on interpretability rather than merely on stress data.

“Physical” Dimensions

In order to interpret the dimensions of the euclidian configuration, multiple regression was carried out with the coordinates of the two dimensional solution as the independent variables and each of a number of dependent variables. Since the goal was to find a mapping between the “physical” space (defined by the parameters) and its corresponding psychological space, the dependent variables consisted of the actual “physical” parameters used and transformations of them:

- E_t: (Parameter E * 100) - 100.
- V_t: (Parameter V * 100) - 100.

---

[4] Both types of scaling led to almost identical configurations.

- R_t: (Parameter R * 100) - 100.
- E_t/V_t
- E_t/R_t
- V_t/R_t
- V_t+R_t: This correspond to Gärdenfors and Holmqvists V’ (see above).
- E_t+R_t: This correspond to Gärdenfors and Holmqvists E’ (see above).
- AbsWidth: An alternative variable (to E_t+R_t) for representing the width. For reasons of simplicity, only the rank order is used here. By definition, stimuli with ($E_t = 10 \wedge R_t = 20$) and ($E_t = 25 \wedge R_t = 10$), got the same rank since they have the same absolute width.
- AbsHeight: An alternative variable (to V_t+R_t) for representing the width. Again, for reasons of simplicity, only the rank order is used here. By definition, stimuli with ($V_t = 10 \wedge R_t = 20$) and ($V_t = 25 \wedge R_t = 10$), got the same rank since they have the same absolute height.
- AbsSum: AbsWidth + AbsHeight.

Many of the independent variables have strong intercorrelation with others. Even so, it is not meaningless to examine them all since it is not given at hand which is best in a set of correlated variables.

Interpreting Configurations

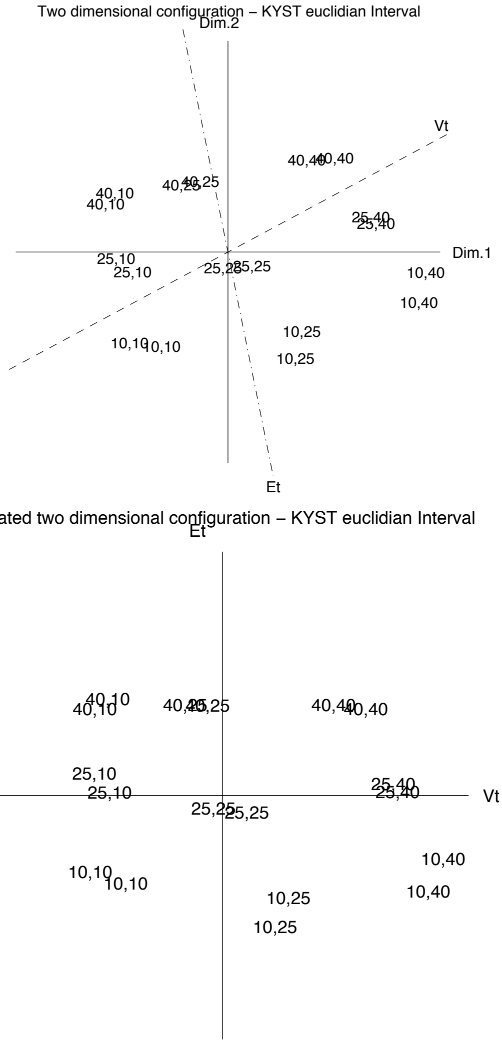
Two Dimensional Configurations

The results of the multiple regression analyses for the dependent variables in the euclidian configuration are shown in Table 1. below. Normalized regression weights (Norm. regr. weights) are the regression coefficients normalized so that their sum of squares equals 1 for every scale. Normalized regression weights could be used as direction cosines.

KYST E	**Norm. regr. weights**		**Mult**		
Variable	**Dim. 1**	**Dim. 2**	**R^2**	**F-ratio**	**P**
E_t	-0,197	0,980	0,955	190,826	0,000
V_t	0,880	0,475	0,962	239,501	0,000
R_t	0,104	0,995	0,008	0,062	0,940
E_t/V_t	-0,764	0,645	0,760	22,594	0,000
E_t/R_t	-0,234	0,972	0,606	11,983	0,001
V_t/R_t	0,898	0,440	0,538	8,913	0,003
$V_t + R_t$	0,875	0,484	0,885	62,952	0,000
$E_t + R_t$	-0,174	0,985	0,837	38,809	0,000
AbsWidth	-0,173	0,985	0,806	31,261	0,000
AbsHeigh	0,859	0,512	0,885	61,428	0,000
AbsSum	0,308	0,951	0,785	28,178	0,000

Table 1. Results from multiple regression analysis for the two dimensional configuration for mollusc shells KYST - euclidian.

The results clearly shows that V_t and E_t (marked with bold text) seems to be relevant dimensions. The determination coefficient is in the 0.90’s, and the multiple correlation is statistically significant at the 0.001 level. It does not seem to be any more relevant “physical” dimensions involved since the other dimensions with high determination coefficients are highly correlated with either of V_t and E_t.


Two dimensional configuration – KYST euclidian Interval

Rotated two dimensional configuration – KYST euclidian Interval

Figure 10. Two dimensional euclidian configuration provided by KYST. Top: unrotated configuration. Bottom: rotated configuration.

Topmost in Figure 10. above, the “physical” dimensions are plotted in the euclidian configuration as provided by KYST. The original axes in the configuration is marked with “Dim.1” and “Dim.2” respectively. The “physical” dimensions are marked with the name of the dependent variable.

The ideal is that the configuration is completely orthogonal, since perpendicular axes are simpler and hence scientifically preferable. The dimensions in the top figure above are not perfectly perpendicular, but note that a slight rotation of V_t clockwise does not draw away to much from the ideal configuration w.r.t. the regression analysis. The observant reader probably have noted that the configuration suddenly “have become” orthogonal (bottommost in Figure .). The reason for this is due to the fact that the author forced V_t to be perpendicular to E_t. As a consequence, determination coefficients etc. from Table 1. should be interpreted with care for this configuration. The two dimensional city-block configuration were analysed with Spearman rank order correlations (for reasons explained above) for each of the dependent variables (see Appendix A). The result were comparable to the euclidian case: V_t and E_t had the strongest correlations (.944 and -.938 respectively) and so were the relevant dimensions also for this configuration.

The determination coefficients for the euclidian configuration are not directly comparable with the correlational coefficients from the city-block configuration. In order to compare them, the Spearman correlation coefficients was calculated also for the euclidian configuration.

The euclidian configuration had slightly larger correlation coefficients (Table 2.) and slightly lower stress values (see Figure 9.) compared to the city-block. So, if these measures are considered alone, the euclidian configuration is preferable, but with a small marginal.

Spearman corr.	**KYST configurations**	
Variable	**Euclidian**	**City-Block**
E_t	0,944	-0,938
V_t	0,944	0,944

Table 2. Comparison of correlation coefficients between the euclidian and the city-block configurations.

To sum up this subsection, even though it is not perfectly clear which configuration and which metric is the “best”, it is clear that a two dimensional configuration is interpretable.

Three Dimensional Configurations

The results of the multiple regression analyses for the three dimensional euclidian configuration are shown in Table 8. below.

V_t and E_t (bold text) still seem to be relevant dimensions. The aim now is to find the third dimension. According to Kruskal&Wish’s criteria for a good interpretation, the candidate variables for the third dimension are marked with italics. However, no candidate is a good candidate for several reasons. One is that they are all strongly correlated with either V_t or E_t. Further, no one of them have a large absolute value for the direction cosine for the third dimension, indicating that the angle between the dimension and the direction of the associated scale is small. This means that no variable from this set could be used as an explanation for the third dimension.

KYST E	**Norm. regr. weights**			**Mult**		
Variable	**Dim. 1**	**Dim. 2**	**Dim 3**	**R^2**	**F-ratio**	**P**
E_t	0,203	-0,976	-0,082	0,976	187,819	0,000
V_t	-0,889	-0,436	-0,142	0,970	149,175	0,000
R_t	-0,101	-0,256	-0,961	0,061	0,303	0,823
E_t/V_t	0,654	-0,512	0,557	0,835	23,674	0,000
E_t/R_t	0,231	-0,967	0,112	0,600	7,010	0,004
V_t/R_t	-0,929	-0,369	0,018	0,542	5,533	0,010
$V_t + R_t$	-0,843	-0,449	-0,297	0,900	41,834	0,000
$E_t + R_t$	0,181	-0,963	-0,198	0,897	40,428	0,000
AbsWidt	0,181	-0,969	-0,169	0,858	28,097	0,000
AbsHeigh	-0,875	-0,483	-0,029	0,873	32,199	0,000
AbsSum	-0,317	-0,939	-0,136	0,808	19,665	0,000

Table 3. Results from multiple regression analysis for the three dimensional configuration for mollusc shells provided by KYST euclidian.

The same anomalies apply for the KYST-city-block configuration: the correlation coefficients for V_t and E_t are high, but there are no high correlation coefficients for any of the dependent variables in the third dimension (see Appendix A). Visual inspection of the configurations gave no hint for some meaningful interpretation either. In other words, a two-dimensional configuration is obviously better for the mollusc shells in this case. Note that this is perfectly in line with the heuristic principle of the “elbow”.

It is interesting to note that one of the parameters used, R_t, seem to have very little importance in the obtained psychological space. R_t alone was actually very badly correlated in both the two- and the three-dimensional solutions. This clearly show that MDS is advantageous over methods that presumes the dimensions involved, even though it could not be completely excluded that R_t had little significance due to the fact that it was only varied between two values.

To sum up this subsection, even though it is not perfectly clear which configuration and which metric is the “best”, it is clear that a two dimensional configuration is interpretable, and therefore preferable, over a three dimensional.

THE BEETLE STUDY

Stimuli

As mentioned, the mollusc shell stimuli have a rather high degree of ecological validity. A problem, though, is that the “physical” dimensions (or parameters) are, at least to some extent, dependent on each other, meaning that the underlying “physical” space is non-orthogonal. The parameter R, the expansion rate of the aperture of the shell, affects both the resulting vertical and horizontal size of the shell. The effect of this nonindependence is that the “physical” dimensions not can be modified independently of each other.

Because of the anomalies with the mollusc shell shapes, a new type of stimuli was designed. The choice fell on beetles. Most people have experience with them, but it is rare that non-entomologists know anything about categorizing them. According to my own experience, the one type of beetle most people know is the ladybird. Besides, any people does not even know it is a beetle. The advantages with mollusc shells are also fulfilled by beetles.

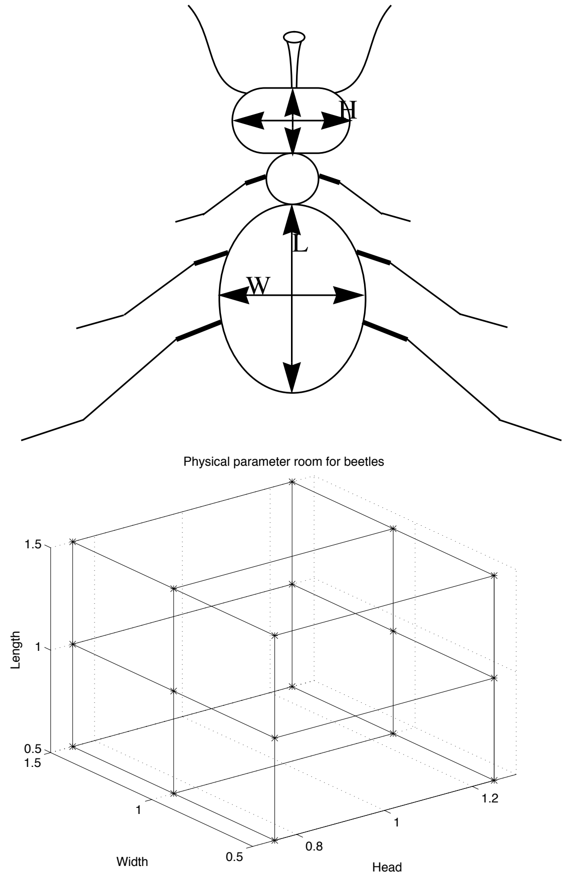


Figure 11. Top: “physical” dimensions for the beetle stimuli. H correspond to the absolute size of the head, L correspond to the length of the abdomen, and W correspond to the width of the abdomen.[5] Bottom: The complete “physical” room for the beetles used in the study.

The beetles used in the following study was created by Niklas Mellegård at University College of Skövde, who did the artistic work, and the author, who stood for the entomological details. The stimuli bears, on purpose, no resemblance to any particular type of existing beetle. Even though, my personal opinion is that the stimuli look realistic. In some cases, beetles bears some resemblance to ants, and in order to decrease this possibly disturbing factor, we provided the beetles with a salient proboscis that was held constant among them. Once a prototype[6] beetle was created, it was easy to create different variations of it by using morphing techniques. The beetle stimuli varied in three “physical” dimensions: the absolute size of the head, the length of the abdomen and the width of the abdomen (see Figure 11.). The head varied between two parameter values, 0.75 and 1.25. The length and width of the abdomen varied between three parameter values, 0.5, 1.0 and 1.5. The parameter values of the head are interpreted as (Parameter value * 100)% area of the head compared to the prototype. The parameter values for length and width have the same interpretation, except that they are relative to the prototype’s length and width respectively rather than to the area. The images corresponding to the complete variation of the “physical” parameters are shown in Figure 12. and Figure 13.

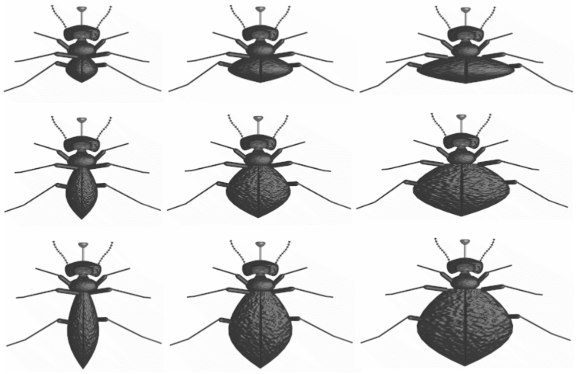


Figure 12. Beetles with the head parameter set to 0.75. The width parameter correspond to the horizontal axis; beetles have, in left to right order, values 0.5, 1.0 and 1.5. The length parameter correspond to the vertical axis; beetles have, in bottom to top order, values 0.5, 1.0 and 1.5.

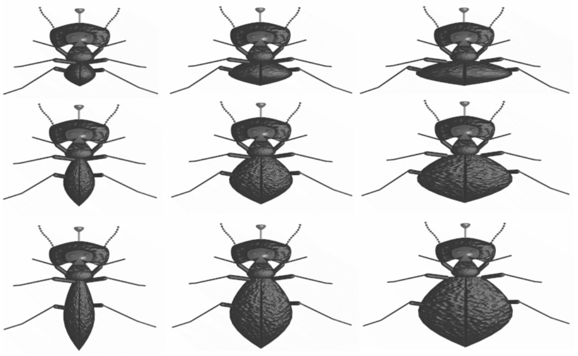


Figure 13. Beetles with the same width and length parameter values as in the figure above. The head parameter is held constant at 1.25.

The proboscis, the antennas, the legs and the thorax is constant among stimuli.

---

[5] This picture is a slight variation of the one presented in (Mellegård, 1995).

[6] The word prototype is meant to be interpreted as in the context of engineering rather than psychology.

Method

Also this study was in form of a session with PsyScope, programmed by Måns Holgersson and the author. The session consisted of a background phase, a presentation phase, an instruction phase and a similarity rating phase.

Before a session the experimenter informed the subject personally that a typical session would take about 20–30 minutes, that there would be breaks and that instructions would be given on the screen.

The Background Phase

The first screen told subjects that the purpose of the study was to investigate how people judge similarity between stimuli, in this case about insects. In contrast to the mollusc shell study, subjects were also provided with some background to the test, in form of a scenario. The scenario told that, some months ago, a zoological expedition arrived at a small island outside New Guinea. There they found a genus of insects that was previously unknown to science. This genus had a characteristic in form of the strange proboscis.

After this, subjects were told that they were to see a collection of pictures of insects belonging to the genus in order to get an idea about the variation of stimuli.

The Presentation Phase

During the presentation phase beetles were presented one by one in randomised order between subjects. Each picture was presented for three seconds.

The Instruction Phase

After the presentation phase, subjects were told that the insects not yet had been divided into species, but that experience says that appearance usually gives a good guidance to this work. Therefore, the first step was to divide them in to groups after their appearance. In order to do this, however, the zoologists needed information about how people judge the similarity/dissimilarity between insects. Subjects were told that this was their specific task.

Further, subjects were instructed to tell how similar or dissimilar the two insects were in each pair they were to see. The ratings should be done using a nine-graded scale where the most similar pair/pairs should have the value “1”, and the most dissimilar pair/pairs should have the value “9”. As in the mollusc shell study, subjects were told that they could base their personal ratings on whatever attribute they felt were relevant, but that they should try to be as consistent as possible during the session.

The Similarity Rating Phase

The similarity rating phase was exactly the same as in the mollusc shell study except for the stimuli. In other words, the same screen layout (see Figure 14.), the same randomisation technique and grouping was used. Also the number of stimuli was the same.

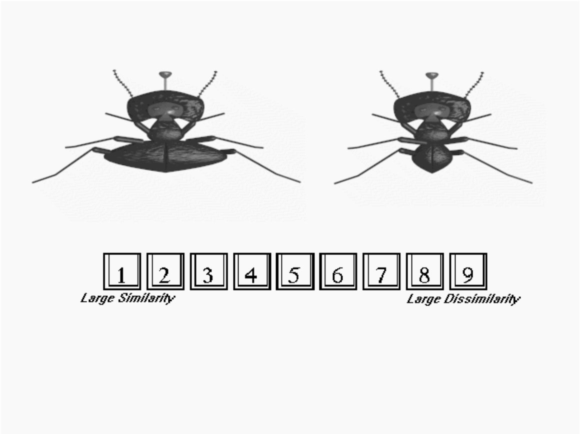


Figure 14. Example of a screen during the similarity rating phase.

Subjects

Ten subjects (most of them students) participated without credit. Four of them were assigned to Group A and six of them to Group B.

Results and Discussion

The same methodology and approach as for the mollusc shells was used, i.e. non-metric scaling at the ordinal level was used first.

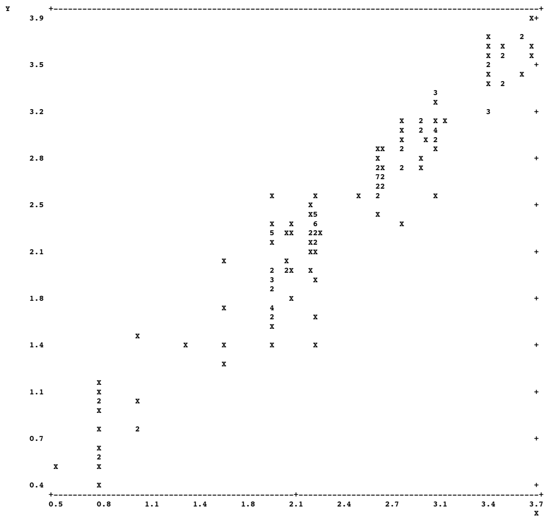


Figure 15. Scatter diagram provided by KYST for beetles: distances (vertical) vs disparities (horizontal)

Since the scatter plot of distances versus disparities (Figure 15.) showed that a linear function could be used, the analysis described is completely based upon scaling at the interval level.

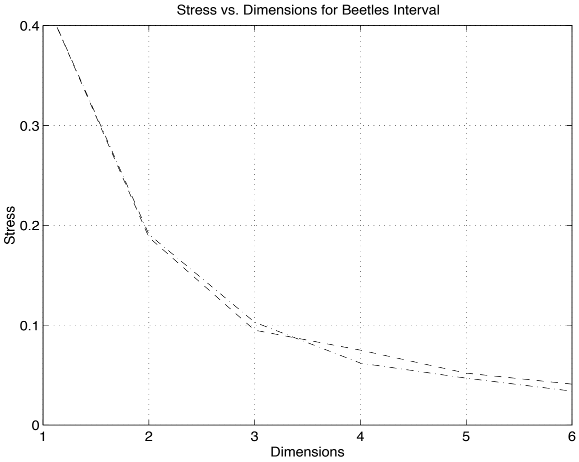


Figure 16. Stress vs. dimensions for beetles using KYST euclidian metric (dash dotted) and KYST city-block metric (dashed).

According to the stress vs. dimensions plot in Figure 16. a three-dimensional configuration should probably be used. However, it might be wise to find out what further explanation we get for the proximities using a three-dimensional solution rather than a two-dimensional. With this question in mind, both the two- and the three-dimensional solutions will be examined.

"Physical" Dimensions for Beetles

The dependent variables used for multiple regression for the configurations was:

- Head: 0.75 for "small" heads and 1.25 for "large" heads.
- Length: 0.5, 1 and 1.5. Note that Length correspond to the relative length of the abdomen.
- Width: 0.5, 1 and 1.5. Note that Length correspond to the relative width of the abdomen.
- Width/Length: The variable correspond to the shape of the abdomen.
- Area: coded as Width * Length. Note that this variable correspond to the relative size of the area of the abdomen.

Interpreting Configurations

Two Dimensional Configurations

The results of the multiple regression analyses (Table 8. below) show that Head and Width/Length seem to be the most relevant variables. The determination coefficients are sufficiently high and the levels of significance are better than .001. Further, Head and Width/Length are approximately perpendicular to each other (Figure 17.).

KYST E	**Norm. regr. weights**		**Multiple**		
Variable	**Dim. 1**	**Dim. 2**	**R^2**	**F-ratio**	**P**
Head	0,936	-0,353	0,878	53,960	0,000
Length	-0,530	-0,848	0,682	16,098	0,000
Width	-0,080	0,997	0,400	4,991	0,022
Wi/Le	0,302	0,953	0,875	52,715	0,000
Area	-0,994	-0,110	0,123	1,050	0,374

Table 4. Results from multiple regression analysis for the two dimensional configuration for beetles provided by KYST euclidian.

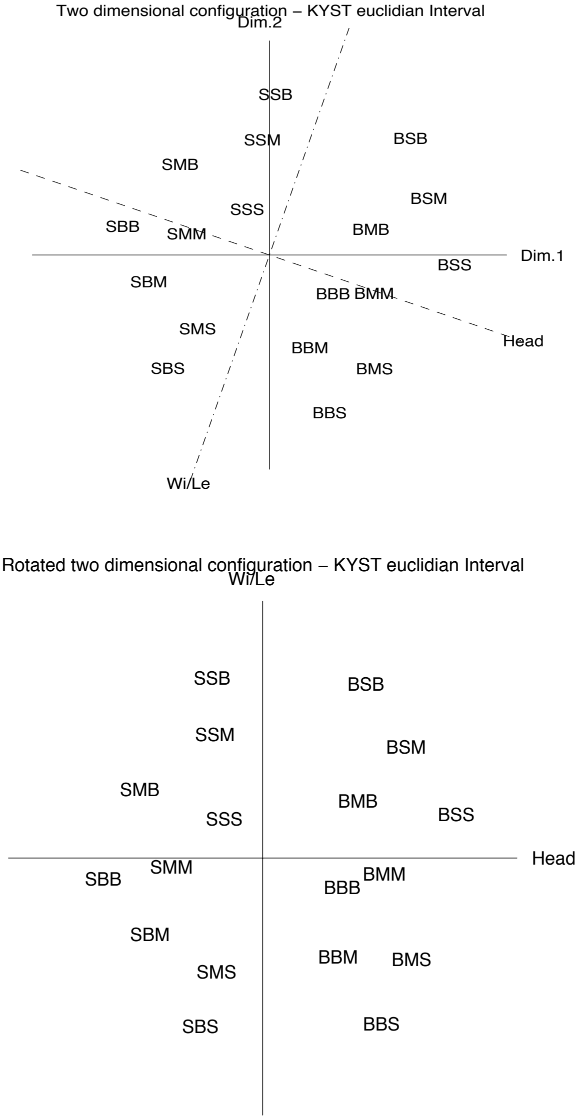


Figure 17. Top: Dimensions Head and Width/Length plotted in the KYST-euclidian configuration. Bottom: Rotated configuration. Stimuli are coded according to their "physical" parameters: Head (S for small or B for big), Length (S, M for medium or B) and Width (S, M or B).

The analysis of the KYST-city-block configuration lead to approximately similar results. Head and Width/

Length have the largest correlation coefficients (see Appendix A), -.867 and .898 respectively.

Goodness of fit are very similar for both configurations, but in terms of Spearman rank order correlation (Table 8.), the euclidian configuration is preferable over the city-block.

Spearman corr.	**KYST configurations**	
Variable	**Euclidian**	**City-block**
Head	0,867	-0,867
Length	-0,787	-0,525
Width	0,577	0,734
Wi/Le	0,979	0,898
Area	-0,479	0,294

Table 5. Comparison of the correlations of the euclidian and the city-block configurations.

Three Dimensional Configurations

Table 8. show the results for the multiple regression analyses for the euclidian three dimensional solution.

KYST E	**Norm. regr. weights**			**Mult**		
Variable	**Dim. 1**	**Dim. 2**	**Dim 3**	**R^2**	**F-ratio**	**P**
Head	0,821	-0,276	-0,499	0,976	188,719	0,000
Length	-0,441	-0,668	-0,600	0,887	36,712	0,000
Width	-0,104	0,615	-0,781	0,881	34,688	0,000
Wi/Le	0,276	0,958	0,076	0,895	39,691	0,000
Area	-0,336	-0,008	-0,942	0,866	30,190	0,000

Table 6. Results from multiple regression analysis for the three dimensional configuration for mollusc shells provided by KYST euclidian.

The three-dimensional configuration is not as simple to interpret as the two-dimensional since all determination coefficients are relatively high. Two sets of dimensions are possible: {Head, Length, Width} and {Head, Width/Length, Area}. Both are approximately equally good w.r.t. the determination coefficients. However, after visualisation of the two alternatives (Figure 18. and Figure 19.), it is clear that the second is the better in that the axes are relatively perpendicular to each other. Also, for example, the obvious order of the beetles w.r.t. the shape of the abdomen is left unexplained if Width and Length are used as primitive axes.

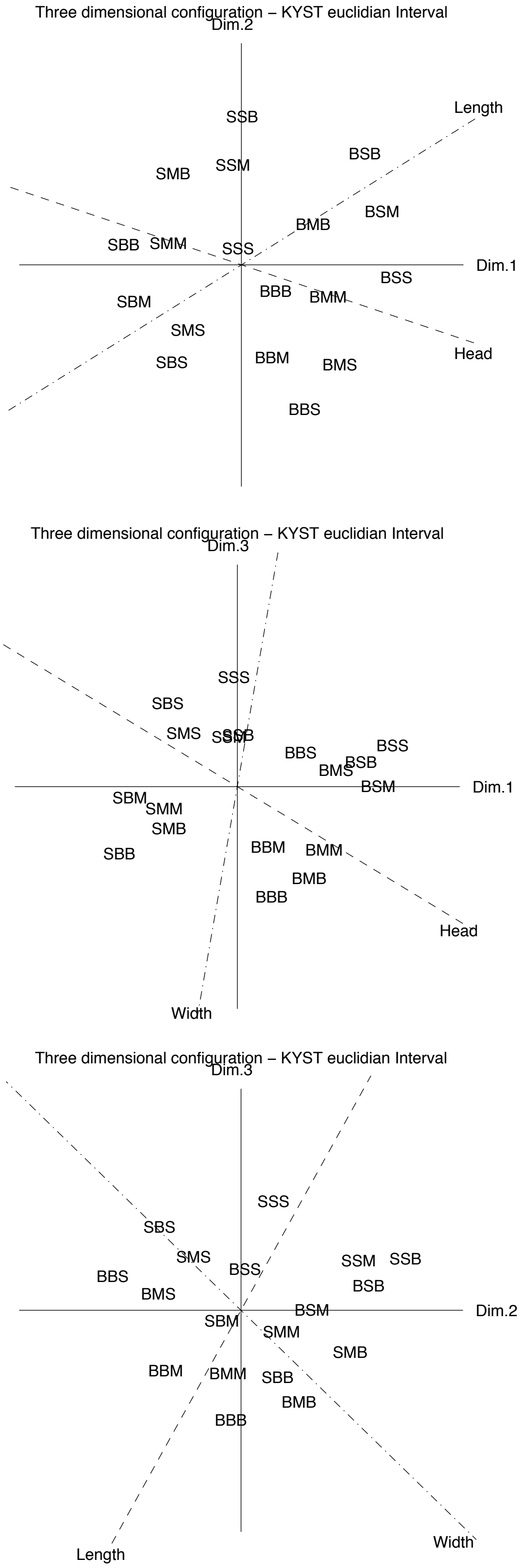


Figure 18. The three-dimensional KYST configuration unrotated. In order from top to bottom, plots of: Head vs. Length, Head vs. Width and Length vs. Width.

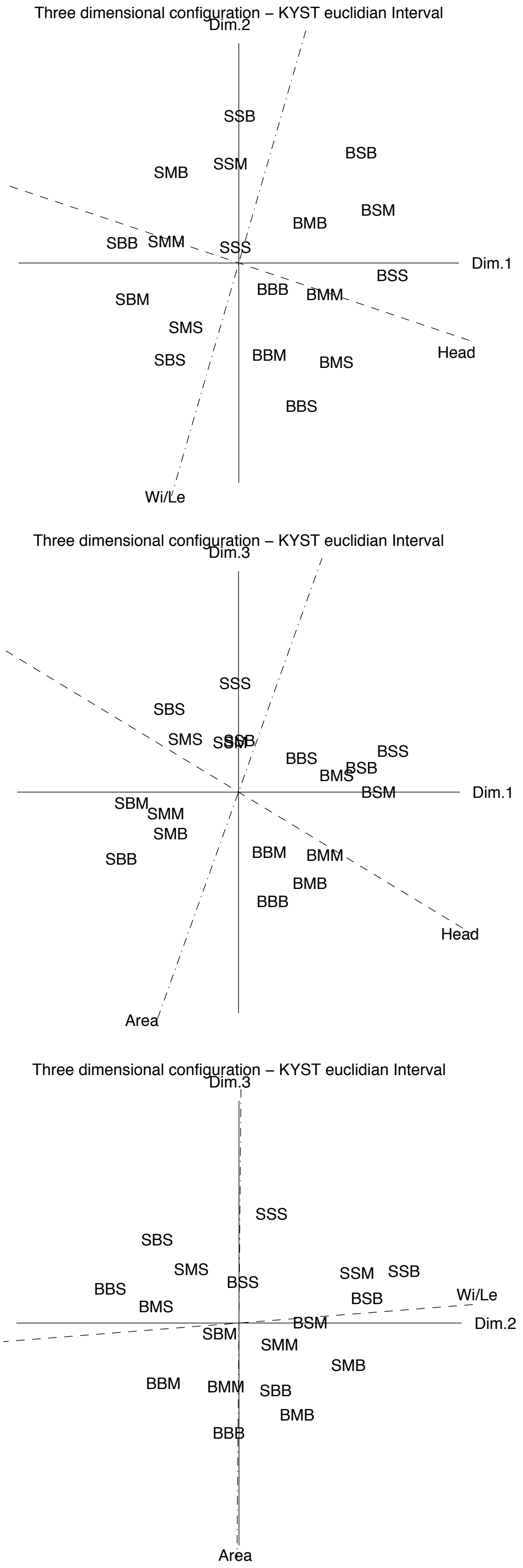


Figure 19. The three-dimensional KYST configuration unrotated. In order from top to bottom, plots of: Head vs. Width/Length, Head vs. Area and Width/Length vs. Area.

Table 8. show the Spearman rank order correlations for the KYST city-block configuration.

KYST C	Spearman corr. coeff.		
Variable	Dim. 1	Dim. 2	Dim 3
Head	0,803	0,503	0,246
Length	-0,551	0,459	0,538
Width	0,092	-0,695	0,656
Wi/Le	0,451	-0,814	0,067
Area	-0,321	-0,158	0,878

Table 7. Results from multiple regression analysis for the three dimensional configuration for beetles provided by KYST city-block.

In this case, merely by inspecting the table of correlations, it is clear that {Head, Width/Length, Area} is a better alternative than {Head, Length, Width}. The differences in correlation coefficients between the sets are substantial.

Of the two euclidian alternatives {Head, Width/Length, Area} is chosen. Before the rotation, Head and Area were to be perpendicular to Width/Length. Since the original configuration was almost orthogonal, only small adjustments was needed (i.e. the axes chosen are almost optimal, and the configuration is unchanged).

The two final three dimensional configurations (i.e. the euclidian and the city-block) could now be compared (Table 8., Figure 20. and Figure 21. below).

Spearman corr.	KYST configurations	
Variable	Euclidian	City-block
Head	0,867	0,803
Length	-0,813	-0,551
Width	-0,559	-0,695
Wi/Le	0,970	-0,891
Area	-0,842	0,878

Table 8. Comparison of the correlations between the euclidian and the city-block configurations.

In both configurations Width/Length seems to be a more important dimension than compared to Area. Especially in the city-block configuration, Head also is weighted heavier than Area.

To summarize this pilot study, the psychologically motivated space underlying similarity judgements between the beetles (at least for this subset) seems to be adequately described by the dimensions Head, Width/Length and Area. With respect to the correlation coefficients alone, the distances between beetles in the space are better described by euclidian distances.

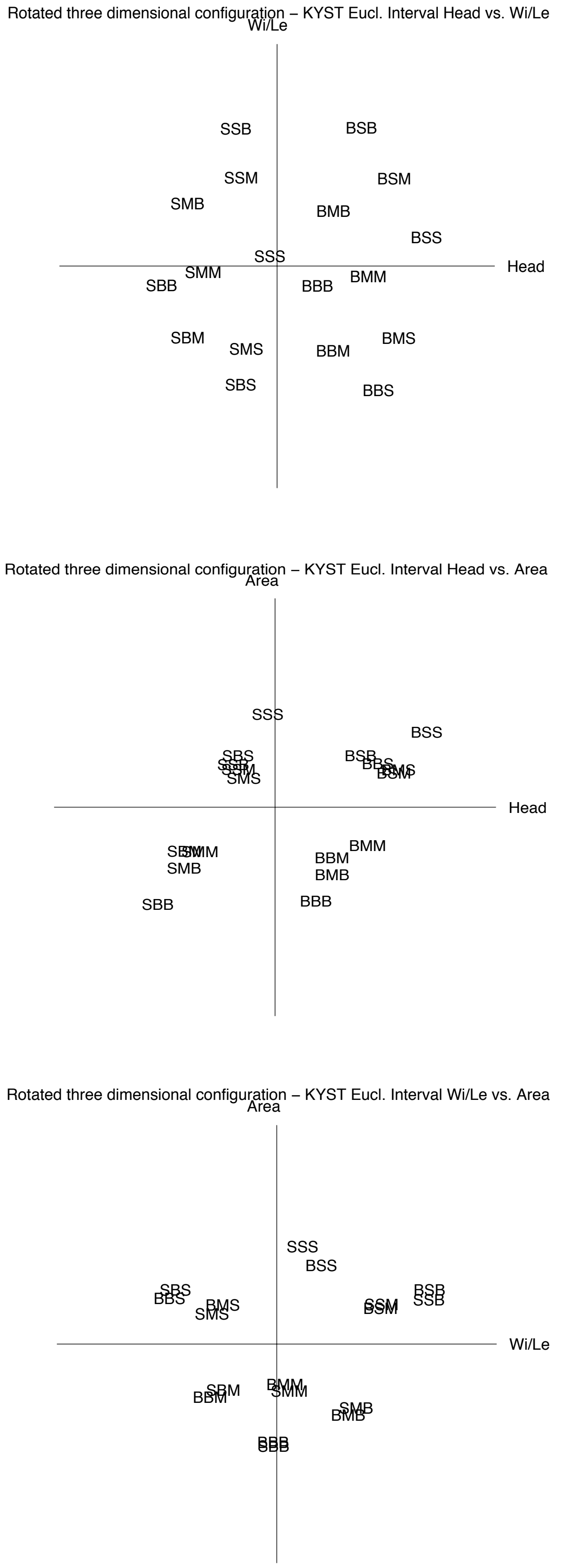


Figure 20. The three-dimensional KYST Euclidian configuration rotated. In order from top to bottom: Head vs. Width/Length, Head vs. Area and Width/Length vs. Area.

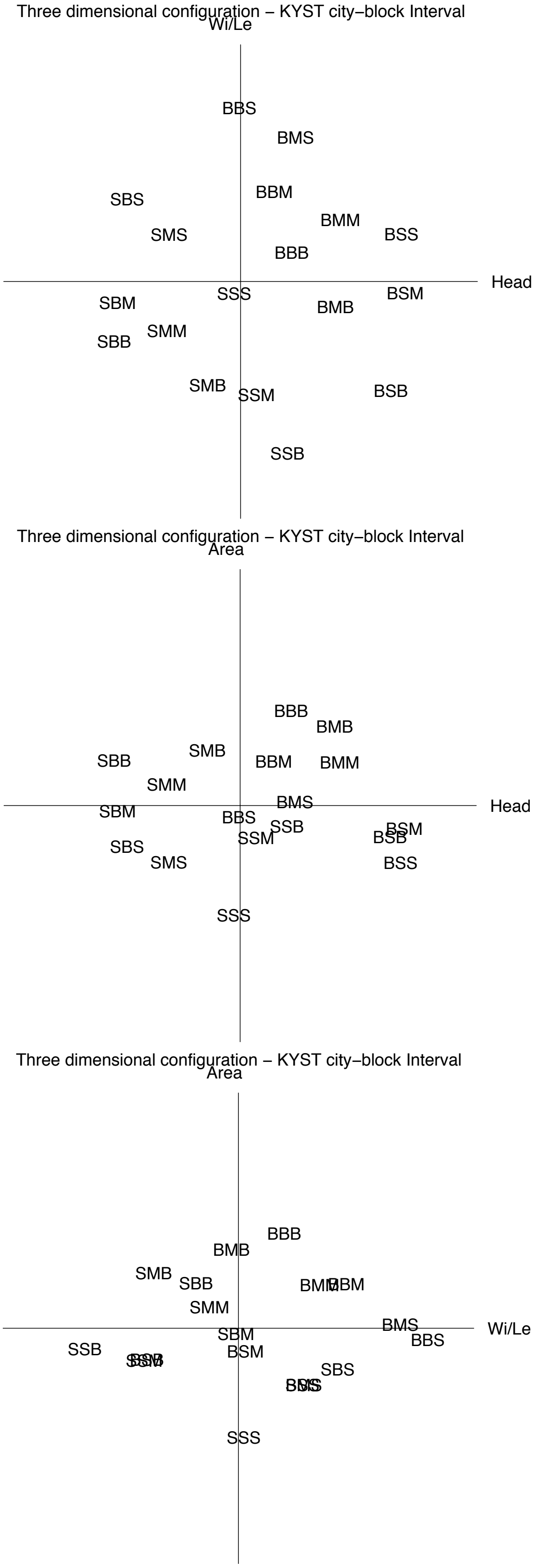


Figure 21. The three-dimensional KYST City-block configuration. In order from top to bottom: Head vs. Width/Length, Head vs. Area and Width/Length vs. Area

GENERAL DISCUSSION

The two studies accounted for both serve as examples of that MDS really is a useful tool for finding a configuration reflecting subjects perceived similarities between stimuli. However, if the configurations will be used as a base for a more general psychologically motivated space, the work does not stop here. There are still some questions that need to be answered.

The Euclidian or the City-Block metric?

First, in the studies above, both euclidian and city-block configurations were derived and compared briefly. It is clear that the configurations derived in the pilot studies were interpretable. However, it was not very obvious which underlying metric described similarities between stimuli the more adequately. The question is on what basis such a conclusion could be derived.

When the choice of the metric is based on the relation between primitive physical parameters and the configuration as above, configurations based the city-block metric has an obvious disadvantage in that they are non-rotatable. However, this problem may be overcome by arranging the points so that they are organised according to the dependent variables and the difference between the stress values of the “original” and the “rotated” configurations are as small as possible.

There are two main classes of stimulus dimensions. Garner (Garner, 1974) call pairs of dimensions that are processed independently (e.g. hue and shape) *separable* dimensions, whereas pairs that are processed as an unanalyzable whole (e.g. hue and brightness) are called *integral*. An operational test for classifying whether stimulus dimensions are integral or separable is to analyse direct scaling data with an MDS-procedure, as described above (see e.g. (Maddox, 1992) for a description of, and especially criticism of, this test). If the euclidian distance metric fits the data best, the dimensions are assumed to be integral. If the city-block metric fits best, the dimensions are assumed to be separable. On the contrary, according to Melara (Melara, 1992), one can almost always find evidence for a city-block distance function for any pairs of perceptual dimensions.

In both pilot studies described above, especially in the case of the mollusc shells, the information obtained does not really suggest which metric is the more suitable. It may also be questioned if this choice should be based on non-substantial differences. However, the choice could be based upon further empirical studies. Under the assumption that integral dimensions are better described by euclidian distances, and that separable dimensions are better described by city-block, some other method, for example a filtering task (see e.g. (Maddox, 1992)), could be used in order to find out whether V_t and E_t are separable or not. In light of the result, the suitable metric and configuration could be used.

Deriving and Using the Psychologically Motivated Space

Now, what can be said about the derived spaces, and how can we use them? First, it is important to be aware of that MDS merely is a tool, or a guide, for understanding. Shepard, discussing MDS (Shepard, 1972, pp. 10–11) warns: “In particular, it may serve as a guide – but never as a substitute – for careful understanding or creative thought.” In other words, care should be taken, both with the search for and with the use of MDS derived configurations.

The configurations derived as above could not be used for evaluating categorization rules etc. directly. They must first be normalized. It is probably not meaningful to “overinterpret” a configuration derived by MDS as above, but rather derive an averaged space reflecting the main characteristics of it. This could be done by deriving mapping functions, either by creative thinking, or with the help of some mathematical tool, e.g. a neural network. Of course, it is always possible to calculate the stress value, even for a configuration derived by the use of mapping functions.

With help of “true” mapping functions and the “physical” parameters for a stimuli, the location in the psychologically motivated space could be calculated. This in turn leads to that categorization rules etc. could be evaluated.

Criticism of Results and Methodology

The studies and the analyses of them may be criticised in several ways. All scalings have been based on averaged ratings from different subjects. The reason for doing so have been to show the methodology, and therefore individual differences have been of secondary importance. However, without taking individual differences into account may result in a psychological space which is not relevant for anyone.

Shepards’ comment from above, that formal tools should not substitute creative thought, was in the context of MDS, but is here believed to be more general. However, it can be questioned how much mathematics, statistics etc. can be violated without the reasoning being *ad hoc*. Some violations must be allowed, otherwise formal tools could certainly be used as a substitute for creative thought. For example, mapping functions derived as suggested above will probably rely upon some assumptions, and the question is if these are relevant or not. In principle, this and other questions regarding the reliability of a psychologically motivated space, should probably be answered empirically. What could be done is to do multidimensional scaling for other sets of stimuli and compare/calibrate the space until no calibration is needed. Certainly, the more observations, the more reliable will the psychological space and the mapping functions be.

SUMMARY

The main purpose with this paper was to describe how a psychologically motivated conceptual space could be obtained with MDS (multidimensional scaling) and how it could be expressed it in terms of a more primitive "physical" (or mathematical) one.

A conceptual space consist of a number of quality dimensions, each representing some quality of a concept. For theoretical reasons especially, it is desirable that a conceptual space is psychologically motivated, i.e. psychological similarity should be reflected by distance relationships between "points" representing stimuli, concepts etc. For practical purposes it is also desirable that the space is expressed in terms of a more primitive "physical" (or mathematical) one. If these requirements are fulfilled, a conceptual space could be used as a base for description and prediction etc. of complex cognitive behaviour like concept formation.

By letting subjects rate the similarity between pairs of systematically varied stimuli and then analyse the ratings with MDS, specific psychologically motivated configurations can be derived. Then, by interpreting and generalising the configurations in terms of "physical" or mathematical parameters, it is possible to derive more general psychologically motivated spaces.

The idea was demonstrated practically with the aid of two relatively equal experimental pilot studies: In the first study, rather "ecologically valid" depictions of mollusc shells were used as stimuli. The stimuli were generated, hence also described, by three "physical" parameters. The MDS analysis indicated that the psychologically motivated space mainly could be described by two dimensions. By testing the different parameters and transformations of these, it became clear that two of the primitive parameters could be used as constituting the dimensions. In the second study depictions of beetles, also these "ecologically valid", were used as stimuli. In this case, the psychologically motivated space came to consist of three dimensions that could be described by one primitive and two transformations of primitive parameters. In both the pilot studies, configurations based on both the euclidian and the city-block metric were analysed. This analysis gave no obvious results, especially not in the case of the mollusc shells.

Possible ways of deciding what metric should be used, like basing the choice upon the results of operational tests of integrality/separability, and possible ways of obtaining a more general psychologically motivated space, were discussed. The paper was concluded by a critical discussion of the method used.

ACKNOWLEDGEMENTS

Research for this paper has been supported by the Department of Computer Science, University College of Skövde. Amongst others, I would like to thank my supervisor Professor Peter Gärdenfors at Lund University Cognitive Science and Dr. Lars Niklasson at University College of Skövde for valuable comments on an earlier draft.

REFERENCES

Chatfield, C. & Collins, A. J. (1980): "Multidimensional scaling", chapter 10 in *Introduction to Multivariate Analysis*, University Press, Cambridge, pp. 189–234.

Cohen, J. D. MacWhinney, B. Flatt, M. & Provost, J. (1993): "PsyScope: A new graphic interactive environment for designing psychology experiments", *Behavioral Research Methods*, Instruments & Computers, 25(2), pp. 257–271.

Garner, W. R. (1974): *The processing of information and structure*, New York: Wiley.

Gärdenfors, P. (1992): "A geometric model of concept formation", in *Information Modelling and Knowledge Bases III*, S. Ohsuga et al. (eds.), IOS Press, Amsterdam, pp. 1–16.

Gärdenfors, P. & Holmqvist, K. (1994): "Concept formation in dimensional spaces", Lund University Cognitive Studies,26, Lund University Cognitive Science, Lund University, Sweden.

Kruskal, J. B. & Wish, M. (1978): *Multidimensional scaling*. Beverly Hills, CA: Sage Publications.

Kruskal, J. B., Young, F. W., & Seery, J. B. (1977): *How to use KYST-2A, A very flexible program to do multidimensional scaling and unfolding*. Murray Hill, NJ: Bell Telephone Laboratories.

Maddox, T. W. (1992): "Perceptual and Decisional Separability", chapter 7 in *Multidimensional models of perception and cognition*, Ashby, G. F. (ed.), Lawrence Erlbaum Associates, Hillsdale, N.J., pp. 147–180.

Melara, R. D. (1992): "The concept of perceptual similarity: from psychophysics to cognitive psychology", *Psychophysical Approaches to Cognition*, Algom, D. (ed.), Elsevier Science Publishers B.V., pp. 303–388.

Mellegård, N. (1995): "The Beetles and the Stones – Learning Dynamics in Multidimensional Conceptual Spaces", MSc-dissertation, Department of Computer Science, University College of Skövde, Skövde, Sweden.

Palmer, S. E. (1978): "Fundamenthal aspects of cognitive representation", in *Cognition and Categorization*, Rosh, E. and Lloyd, B. B. (eds.), Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 259–303.

Rosch, E. (1978): "Principles of categorization", chapter 2 in *Cognition and Categorization*, E. Rosch och B. B. Lloyd (eds.), Lawrence Erlbaum Ass., Hillsdale, NJ, pp. 27–48.

Shepard, R. N. (1962): "The analysis of proximities: Multidimensional scaling with an unknown distance. I and II.", *Psychometrika*, 27, pp. 125–140, pp. 219–246.

Shepard, R. N. (1972): "Introduction to Volume I", in Multidimensional Scaling – Theory and Applications in the Behavioral Sciences, Volume 1 – Theory, Shepard, R. N., Romney, A. K. and Nerlove, S. B. (eds.), Seminar Press

Appendix A

– Spearman rank order correlations for configurations based on the city-block metric

KYST C	**Spearman corr. coeff.**	
Variable	**Dim. 1**	**Dim. 2**
E_t	-0,315	-0,938
V_t	0,944	-0,243
R_t	0,075	-0,011
E_t/V_t	-0,877	-0,502
E_t/R_t	-0,298	-0,794
V_t/R_t	0,749	-0,169
$V_t + R_t$	0,925	-0,235
$E_t + R_t$	-0,279	-0,900
AbsWidth	-0,254	-0,861
AbsHeigh	0,887	-0,236
AbsSum	0,413	-0,735

Table 9. Spearman rank order correlations for the two dimensional configuration for mollusc shells provided by KYST - city-block.

KYST C	**Spearman corr. coeff.**		
Variable	**Dim. 1**	**Dim. 2**	**Dim 3**
E_t	-0,275	-0,944	-0,007
V_t	0,944	-0,262	-0,131
R_t	0,096	-0,032	-0,225
E_t/V_t	-0,865	-0,481	0,143
E_t/R_t	-0,298	-0,793	0,063
V_t/R_t	0,749	-0,185	-0,122
$V_t + R_t$	0,931	-0,260	-0,191
$E_t + R_t$	-0,235	-0,912	-0,072
AbsWidth	-0,207	-0,878	-0,067
AbsHeigh	0,900	-0,254	-0,153
AbsSum	0,459	-0,750	-0,067

Table 10. Results from multiple regression analysis for the three dimensional configuration for mollusc shells provided by KYST city-block.

KYST C	**Spearman corr. coeff.**	
Variable	**Dim. 1**	**Dim. 2**
Head	-0,867	-0,375
Length	0,459	-0,525
Width	-0,039	0,734
Wi/Le	-0,350	0,898
Area	0,294	0,143

Table 11. Results from multiple regression analysis for the two dimensional configuration for beetles provided by KYST city-block.