Doctoral Thesis Presented By

Byongho Choe

Submitted At The University Of Oldenburg

In Order To Obtain The Grade Of
Doctor Of Philosophy
Specialization: Psychological Acoustics

NONMETRIC MULTIDIMENSIONAL SCALING OF COMPLEX SOUNDS

DIMENSIONS OF PREFERENCE RATINGS AND PERCEIVED SIMILARITY OF VEHICLE NOISES

Oldenburg, 10. February 2001

Erster Gutachter	Prof. Dr. August Schick
Zweiter Gutachter	Prof. Dr. Klaus Peter Walcher
Datum der Disputation	8. Juni 2001

ACKNOWLEDGMENTS

"The word "yellow" is necessary because there are yellow things; the word "similar" is necessary because there are pairs of similar things. And the similarity of two things is as truly a nonlinguistic fact as the yellowness of one thing." (from An Inquiry into Meaning and Truth, Betrand Russell, 1940)

I am greatly indebted to Prof. Dr. August Schick with whom I have had the honor of working, for his humane care and his scientific contribution to this research. He led me to the study of human sound evaluation and so the area called "Psychological Acoustics". Until two years ago, I didn't know about it. At the beginning of this time I realized that I was incompetent at working the semantic problems as I studied the content analysis of peoples' free descriptions concerning a set of vehicle noises. My understanding of this subject has been strenghened by years of debate and interaction with Prof. Dr. August Schick. Being exposed to sound, we certainly do rely upon auditory perception contained in human minds, or evoked from the sound of its own. There also is a lot of information out there in the sound, not in the head. And, yes, what about my inability to figure out simple or complex relationships between semantics and acoustics? When dealing with the relation between automotive product sounds and subjective perceptions, as I more and more walked into the multidimensional or multifactorial phenomenon of auditory perception, I realized that there appears to be a fundamental difficulty, say a non-linearity of human perceptual processes, in arriving at a metric that will predict the perceptions of product sound quality, not to mention the definition of "good sound". While identifying, distinguishing, or ordering acoustic objects, like any other objects (e.g. visual, tactile or olfactory ones), we observers take into account some individual's subjective equivalents more or less specific to the objects. I hear sounds from complicated things in my acoustic environment in which I live and work. Why can I judge acoustic objects on which I focus? Why can I make contents of perceived or judged sounds in terms of liking vs. disliking, or similar vs. different, visible? Hence this topic: Nonmetric Multidimensinoal Scaling of Complex Sounds. Dimensions of Preference Ratings and Perceived Similarity of Vehicle Noises. The paper presented here as a promotion for a doctor grade of philosophy was confined to the areas moving between psychological acoustics and applied statistics.

I would like to express my appreciation and indebtedness to Prof. Dr. Klaus Peter Walcher for his humane support and for valuable advice, especially for teaching me the area called "Environmental Psychology", and the colleagues at the Institut zur Erforschung von Mensch-Umwelt-Beziehungen (Institute for Research into Man-Environment-Relations, for short "MUB") for sugge-

stions and the use of a laboratory. First of all, the patient critique and helpful comments of Dr. Wolfgang Ellermeier, Institute of Experimental Psychology, University of Regensburg, on an earlier draft of this paper are gratefully acknowledged. Great thanks are due to the people who took the time to talk on the issues and to read the previous versions of this manuscript with care, giving me critical reviews: Dr. Hans-Peter Bäumer, Department of Applied Statistics, University of Oldenburg, for his helpful advice about NMDS techniques, Dipl.-Inform. Andreas Marshall, Department of Applied Statistics, University of Oldenburg, for his critical evaluation of the draft material and for helping to develop the computer programs for the successful running the experimental sessions, and Dipl.-Phys. Matthias Vormann, Department of Physics, University of Oldenburg, who friendly analyzed the acoustic materials via BAS instrument and for his valuable manuscripts relating to the calculation models of psychoacoustic magnitudes. Appreciation is expressed to Dr. Stefan Seuring, Department of Business Management, University of Oldenburg, and Cara H. Kahl, who considerably improved readability by many suggestions on English style. Also, gratitude is due to Hyunsook Koo & Reza Mohadess, Beate Kurz & Dr. Herbert Kurz, Mareike Seuring & Dr. Stefan Seuring, Renate Wenzel & Andreas Marshall, Dr. Hwajeong Kang-Petermann & Wolfgang Petermann, Eunok Bak & Dr. Byeongcheol Gang, Edith Schneebauer & Adolf Schneebauer, Dagmar Ziegler & Dr. Andreas Brüch, Andrea Walther, Dr. Hyeyoung Song, Wolfgang Nußbaumer, Seonghwan Jeon, and Dr. Sharif Notash, who all supported me emotionally so that I could more easily manage living in Germany. Finally, my greatest thanks go to my father, Yoonsik Choe, my mother, Minza Jun, and my brother, Hyong Choe, who supported my foreign studies over ten years with dedication. I cannot find a word for their love.

CONTENTS

1. Introduction

How can sounds produced by a car be subjectvely qualified, and what are the most relevant auditory aspects of car interior sound quality? Sound quality of a vehicle's interior and/or exterior noise is thought of as one of the major design parameters by which one can discriminate similar products (e.g. see Blauert, 1996; Keiper,1997). In the psychoacoustic community it is widely accepted to assume that perceived sound quality elicited by automotive sounds is realized as multi-dimensional. A variety of verbal expressions used to describe the perceived sounds can be indicative of its multidimensional character (e.g. see Gabrielsson & Lindström, 1985). The term "multidimensionality" refers to the very fact that perceived sound quality is made up of a number of separate perceptual dimensions describing different features of the sound to which a subject is exposed (e.g. Bodden, 1997). Moreover, there is a general agreement among psychoacousticians that the high number of parameters necessary to represent the acoustic waves at the ears of listeners is to be reduced to a handful of sensory attributes of auditory events, namely less than four in form of adjectives or expressions which we actually use when it comes to sound quality rating (Blauert & Jekosch, 1997). However, Schick (1979, 1990a, 1994, 1995, 1996), in surveying much of the work on sound quality assessment, has concluded that individuals may attend to different dimensions in making an overall sound assessment on different sounds. It surely seems that there are not merely to acoustic and/or psychoacoustic metrics reduceable cognitive and/or aesthetic aspects which a human subject attaches optionally to the objects present in our daily environment when assessing the questions of sound quality. Respective of auditory perceptions and auditory impressions of vehicle noises, Schick even speaks of "car semantics" commenting that *"The human being regards all these sensations and emotions as the reality of his everyday existence: they determine this daily actions in a great variety of ways. We describe them as being subjective, because only a single "I" of "self" "determines" (auditory) phenomenon." (Schick, 2000, 2521)*. When pushed a little further, we can pose a question as to whether and what kind of aspects or multi-path components for sound quality assessment are considered relevant whatsoever, which might be more or less identifiable with a series of signal variables. In terms of Fastl (1998), one may question to ask how the semantics of hearing or descriptive terms (e.g. crumpled, muffled, palparable, buzzy etc.) could be "translated" into the acoustics of sound or design guidelines (e.g. level, spectrum etc.). This enables, perhaps, the noise control engineer or sound designer in car manufacturing to identify which sound quality characteristics need to be modified to achieve the desired sound quality that drives acceptance or preferences for a product.

Table I-0: Identified acoustic and psychoacoustic indices with regard to the evaluation of sound quality

Names	Sensory Attributes	Theoretical Backgrounds/Methods
Helmholtz [63], Guernsey [28], Lundin [47], Terhardt [84]	Tonality, Consonance	Theory of Musicological Consonance
Helmholtz [96]	Loudness, Pitch, Timbre	Theory of Sensory Euphony
Hornbostel [26]	Loudness, Pitch, Volume, Density	Bandpass Noise Characters
Stevens [34]	Density, Brightness	Method of Magnitude Estimation and Production
Osgood [52][75], Hofstätter [75], Kastka-Buchta [77]	Valency, Activity, Potency	Semantic Differentials
Plomp-Levelt [65]	Fluctuation Strength	Theory of Sensory Euphony
Hawel [67]	21 Bipolar Descriptors	Factor Analysis using 380 Pairs of Adjectives
Kryter [71]	Loudness, Roughness, Sharpness, Pleasantness	Theory of Sensory Pleasantness
Bismarck [74]	Sharpness, Density, Hardness, Brightness	Semantic Differentials & MDS
Terhardt-Stoll [81]	Fluctuation Strength, Sharpness, Roughness, Tonality	Pairwise Preferences on 17 Environmental Noises
Terhardt [74][81][84], Mellert-Weber [78][81], Zwicker [82], Aures [85], Hellman [85], Preis [87], Weber [90], Kohler-Kotterba [92], Heldman [94], Daniel [95]	Roughness, Sharpness, Fluctuation Strength, Sonorousness	Annoyance on CISQ & Pair Comparison Methods
Gabrielsson-Lindström [85]	Bipolar Adjectives	Sound Quality of Sound Reproduction Systems
Höger [86][87]	Pitch, Brightness, Harmonic Density, Evaluation	Content Analysis & MDS
Boemak [94]	Brightness, Softness	Assessment of Car Engine Noises
Bisping [94]	Sharp-Dull, Metallic-Low	Assessment of Car Door Slam Noises
Bisping [94], Bisping-Giehl [96]	Pitch, Impulsness, Timbre	Assessment of CISQ
Bisping [95][97]	Strong-Weak, Pleasant-Unpleasant	Assessment of Car Engine Noises
van der Auweraer et al. [97]	Solidity	Pairwise Preferences on Car Door Slam Noises
Beidl et al. [97]	Loudness, Sharpness, Periodicity, Impulseness	AVL-Annoyance Index

It is known from the bulk of the psychoacoustic work that sensory pleasantness of sound sources (e.g. Aures, 1984a,c,d; Terhardt & Stoll, 1978, 1981; Terhardt, 1984; Weber, 1990; Ellermeier et al., 1997), or annoyance of sounds (e.g. Berglund, 1981; Cardozo & Van Lieshout, 1981a; Kastka, 1981; Mellert & Weber, 1981; Zwicker, 1991; Widmann, 1992; Beidl & Stucklschwaiger, 1997), is influenced by elementary hearing sensations like loudness, roughness, sharpness, tonality, pitch etc., just to name a few (see

Table I-0). Due to Cardozo and Van der Veen (1979) as well Cardozo and Van Lieshout (1981b) a multidimensional annoyance space can be constructed along with sound level on a high-low scale and "sound character" on a good-bad scale. The concept of sound character (in German "Klangcharakter") that originally goes back to Stumpf (1883) is introduced as an additional physical feature responsible for any systematic differences in annoyance caused by different sounds that have the same A-weighted pressure level. Having evaluated engine noises of cars, Boemak (1994) arrived at two main perceptual dimensions: brightness and softness, demonstrating that the total variance extracted was 62% on brightness and 21% on softness, respectively. In the case of car outdoor noise, Bisping (1994) asked for sensory attributes such as a sharp-dull or a metallic-deep scale. After having caried out a series of experimental studies, Bisping (1995, 1997) summarized that car interior sound quality is made up of two major perceptual factors: pleasantness and powerfulness, accounting for 60-70% of the total variance in standard driving situations. As was already pointed out by several authors (e.g. Blauert, 1986; Bodden, 1997; Genuit, 2000), one reason to explain different psychoacoustic results might be that the definition of "sound quality" is at variance, demonstrating that a sonic quality assessment would mostly be brought about by a variety of physical (e.g. energy), psychoacoustic (e.g. loudness), design (e.g. product color) and cognitive aspects (e.g. lifestyle, fashion).[1]

Since Stumpf's pioneer work, "Tonpsychologie" (Stumpf, 1883, 1890), musicological attempts have been made to characterize the dimensionality of timbre perception (e.g. Hornbostel, 1928; Koehler, 1923; Schole, 1930; Stevens, 1934; de la Motte-Haber, 1972; Jost, 1973; Wellek, 1975; Albersheim, 1975). A general review on the perception of instrumental timbres in psychometric works is given in the contribution of Jost (1967). Multidimensional scaling techniques have been applied to explore the perceptual similarities of vowel spectra (e.g. Plomp et al., 1965; Pols et al., 1969), of musical timbres (e.g. Levelt et al., 1966; Plomp, 1970; Miller & Carterette, 1975; Grey, 1977; Krumhansl, 1989; Kendall & Carterette, 1991; McAdams et al., 1995; Donnadieu & McAdams, 1996; Misdariis et al., 1998; McAdams & Winsberg, 1999, 2000), etc. As yet, considerably little research has focused on the mechanisms or processes that determine the perceptual similarities of vehicle interior noises which represent a relatively unstudied aspect of auditory perception. Inspection of a large number of psychoacoustic reports showed that the application of multidimensional scaling techniques on perceived similarity concerning automotive sounds is relatively poor from an experimental psychological standpoint, not to mention the real psychological

[1] e.g. Guski (1997) discussed definitions of sound quality for which three main aspects should be respected at large: (1) stimulus-response compatibility, (2) pleasantness of sounds, and (3) identification of sounds. Bodden illustrated a general procedure for assessing the sonic quality regarding the identification of product-specific noise components (Bodden, 2000).

nature of different perceptual dimensions related to perceived similarity. In addition, there have been few attempts to reflect on the issue of geometric configurations based on both preference and similarity responses to vehicle interior noises. In recent years, some reports on an extended version of the multidimensional scaling algorithm CLASCAL (Winsberg & De Soete, 1993) for subjective evaluation and objective metric correlation work have been made and are currently underway by the Perception and Cognition team at IRCAM (Insititut de Recherche et de Coordination Acoustique/Musique) centered around McAdams (McAdams et al., 1998; Susini et al., 1999), in collaboration with the automobile manufacturers Renault and PSA Peugeot Citro n.

The research reported here was guided by methodological as well as by substantive questions:

(1) Does the monotony condition that the rankings of metric distances reflect those of dissimilarities among the pairs of stimuli to be met?
(2) What sensory attributes of vehicle interior noises are most important in influencing the judgments of preference and perceived similarity?
(3) How many dimensions at most are necessary for a perfect representation if the proximity data can be mapped into Minkowski distances?
(4) Do the sensory attributes that govern preference ratings also impede the judgments based on perceived similarity?
(5) To what extent do the sensory attributes revealed in both preference and similarity structure have a closer link to certain acoustic or psychoacoustic parameters to be technically measured?
(6) Once the relationship between composite metric distances underlying dimensions of subjective ratings and the sum of metric distances pertaining to component dimensions turns out to be linear, can then the integration of prevailing physical factors be based upon an adding model representing psychological processes concerning the sound evaluation?

This study is directed toward determining the number of psychologically meaningful perceptual dimensions required for pairwise ratings of preference and similarity with respect to a set of 15 vehicle interior noises, and toward identifying the physical bases underlying the perceptual dimensions. Nonmetric scaling analysis of the data sets revealed two dimensions, each of which might be roughly interpreted as strength and rhythm. 2D configurations extracted 75% of the variance of perceptual dimensions pertaining to the preference ratings, 71% of the variance of sensory attributes pertaining to the perceived similarity, respectively. A statistically reliable correspondence was observed between the perceptual dimensions and the physical parameters: The first similarity dimension (48% of the variance) was represented by the acoustic parameter sound level (correlation of 85%), the second similarity

dimension (24% of the variance) by the psychoacoustic parameter sharpness (correlation of 66%). Accordingly, the first preference dimension (66% of the variance) could be accounted for by the acoustic parameter sound level (correlation of 92%). The interrelatedness among the first preference dimension and the first similarity dimension amounted to a correlation of 70%, and the corresponding one between the second similarity dimension and the second preference dimension had a correlation of 29% which was not significant. Sharpness in acum appeared to be less suitable for the description of the second preference dimension (10% of the variance), resulting in a correlation of 35% having no statistical significance.

2. Theoretical Part

We were interested in determining whether the nonmetric multidimensional scaling procedure developed by Kruskal (1964a,b) could be fruitfully applied to the determination of a parsimonious number of perceptual dimensions underlying pairwise preference ratings and pairwise similarity judgments with respect to vehicle interior noises. Of great concern was also to see whether a multidimensional solution for averaged dissimilarity data is to be based physically. To begin, briefly, some geometric models will be sketched. Details of multidimensional scaling techniques and rationales have been reviewed by Ahrens (1974) and by Borg and Groenen (1997), so that only some relevant aspects shall be sketched here. For more than a decade, multidimensional scaling, MDS for short, has been thought of as an explorative multivariate technique and has mainly been applied to capture the number of dimensions required to represent perceptual and/or cognitive attributes of stimulus objects in a low-dimensional, extended Euclidean space. MDS is like a heuristic approach permitting investigation of a stimulus domain without preselection of component dimensions by determining how a subject has cognitively classified the objects. MDS is based on the assumption that similarity is thought of as ‘psychological distance’ (Shepard, 1957, 1960). Following Coombs (1952), a multidimensional set of scales is sought directly from ratings of the relative similarities of pairs of stimuli in which the stimulus objects are viewed as points in some latent psychological space and the ordering of similarities is assumed to reflect the order of their metric distances. The best-fit multidimensional metric space of stimulus objects in which the ordering of metric distances in any one of a wide variety of distance geometries is maximally inversely correlated with the ordering of similarities is called a geometric map of stimulus points. Factors or features that might explain the ordering of stimulus points along the coordinates in a latent psychological space of appropriate dimensionality can be interpreted as

perceptual dimensions or sensory attributes that are characterized through the senses (e.g. Carroll & Wish, 1974a). MDS decomposes the similarity data into a set of mutually orthogonal dimensions or attributes. MDS maps metric distances of stimuli into some multidimensional metric space that has the number of perceptual dimensions to be specified by the investigator. The identification of subjective attributes is post hoc, as in Factor Analysis (Spearman, 1927).

The earliest papers on MDS paid more attention to scalar products than to distances. MDS that starts with metric distances is generally of fewer dimensions and is easier to interpret than FA. The latter starts with scalar products generated from e.g. containment judgments, operating on the basis of Euclidean distances. Despite the fact that the "simple structure" criterion of FA is not necessarily relevant, while Factor Analytic rotation in the coordinate system may yield psychologically interpretable dimensions, Stewart (1974) advocated that the multidimensional representation is contained in the Factor Analytic space, and concluded that FA can provide a firm basis for evaluating the generality of the dimensions derived by MDS. Even though scalar products are easier to handle numerically, the overriding benefit of MDS is that spatial models have its virtue in high predictability (e.g. Kendall & Carterette, 1991). Scalar products and Euclidean distances stand in an inverse monotonic relation to each other, so that if Euclidean distances grow, scalar products get smaller, and vice versa. For formal relations between scalar products and Euclidean distances, see Tucker (1972), Borg and Groenen (1997). In a sense, independent variables in MDS are thought of as internal dimensions with specific proximity values to which metric distances as dependent variables are to be allocated.

MDS is eventually called Similarity Structure Analysis (Borg & Lingoes, 1987), or Smallest Space Analysis (Guttman, 1968). MDS is of two types: (1) metric multidimensional scaling, and (2) nonmetric multidimensional scaling. The advent of classical, metric MDS is credited to Eckart & Young (1936), Young & Householder (1938), Richardson (1938), Torgerson (1952, 1965), and Gower (1966). The foundations of nonmetric MDS (NMDS) are laid down by Shepard (1962, 1966), Kruskal (1964a,b), Guttman (1968), Beals et al. (1968), Tversky & Krantz (1970), and Young (1970). For a general review of nonmetric multidimensional techniques, see Gigerenzer (1977, 1981), which includes a good bibliography. A large number of applications of MDS including factor analytic techniques for dimensional and metric representations of an increasingly broad variety of perceptual or cognitive objects, including colors (e.g. Ekman, 1954; Guttman & Kalish, 1956; Helm & Tucker, 1962; Hyman & Well, 1968; Carroll & Chang, 1970), geometrical figures (e.g. Silver et al., 1966; Krantz & Tversky, 1975), nations (e.g. Wish et al., 1970), drawings (e.g. Berlyne & Ogilvie, 1974; Berlyne, 1975; O'Hare, 1976), faces (e.g. Schlosberg, 1954; Abelson & Sermat, 1962; Cliff & Young, 1968; Milord, 1978), personalities (e.g. Walters & Jackson, 1967; Cliff et al.,

1973; Mueller, 1974), soft drinks (e.g. Cooper, 1973), words (e.g. Osgood, 1962; Cliff, 1968; Wainer & Kaye, 1974), attitudes (e.g. Abelson, 1955; Messick, 1956; Shikiar, 1974), inkblots (e.g. Ahrens, 1970), tastes (e.g. Russell & Gregson, 1968; Schiffman & Dackis, 1975), morse codes (e.g. Shepard, 1963), speech sounds (e.g. Shepard, 1972; Howard & Silverman, 1976), etc., all of which would attest to their popularity and usefulness.

2.1 Spatial Representation of Similarity

What is it like to make things appear alike or different? The characteristics with respect to which things are similar may be conceptualized either as common elements (e.g. Ekman, 1954, 1965), or as dimensions on which things have some degree of proximity. According to the data theory improved by Coombs (1960, pp. 144 as well as 158), the judgment relation is either an order relation (comparison of dominance) when the stimuli come singlely or a proximity relation (comparison of consonance) when the stimuli come in pairs. Proximity thereby is thought of as a measure for how two things are alike or different to each other. Judgments of preference or similarity require rather optional performance than optimal performance. Similarity judgments provide information about the ways the stimuli are perceived or recognized (e.g. Tversky, 1968). The concept of psychological similarity is widely employed throughout psychology as descriptive or explanatory concept in those areas: pattern recognition (e.g. Wertheimer, 1923; Olson & Attneave, 1970), learning theory (e.g. Skaggs, 1925; Osgood, 1949; Künnapas, 1966; Miller, 1969; Nelson & Nelson, 1970), biophysics (e.g. Landahl, 1945; McCulloch & Pitts, 1947; Hebb, 1949). Reviews of this literature related to the definition of psychological similarity and the similarity models can be found in papers by Gregson (1975, 1976). Carnap (1967) took the view that the prior requirement for measurement of similarity is an ordering relation on an unidimensional continuum, and proposed a sweeping definition of similarity, putting "*A relation is called a similarity relation if it is symmetrical and reflexive, and an equivalence if it is also transitive.*". His set-theoretical model of similarity has been essentially extended by Tversky & Krantz (1970) and by Tversky (1977). The algebraic foundations of spatial models of similarity have been reviewed by Shepard (1974) and by Gigerenzer (1977).

2.1.1 Minkowski Power Metrics

The object of MDS is to reveal the relationship among stimuli by representing them in an extended Euclidean space of low dimensionality in such a manner

that the ranking order of the pairs of stimuli reflects the ranking order of their metric distances. As in Torgerson's MDS model (1958), metric distances retained from the similarity judgments among the pairs of stimulus objects are factor analyzed (cf. Coombs, 1958). To do this, the similarity data is modeled by one of a family of Minkowski models used to represent metric distances among the pairs of stimuli. The first workable algorithms for determining the metric parameter in Minkowski power metrics are those due to Shepard (1957, 1962), Kruskal (1964a), and later Lingoes and Roskam (1973). The classical MDS is implemented in such a MDS algorithm, MDSCAL, for ordinal scaled data, which plays a central role throughout this paper to be reported (e.g. for an alternative algorithm TORSCA see Young & Torgerson, 1967). The Minkowski distance formula is taken as a theory of how a dissimilarity judgment on two stimuli is generated. The similarity data are converted into metric distances between stimuli i and j, which can be formulated by the Minkowski distance function (Kruskal, 1964b, 116):

$$d_{ij} = \left[\sum_{k=1}^{m} |x_{ik} - x_{jk}|^r \right]^{1/r} \tag{1.1}$$

where x_{ik} is the coordinate of stimulus i on the dimension k. The distance function, i.e. a metric, which describes the "rule of combination" (Shepard, 1964), is a scale that assigns to every pair of points i and j a real nonnegative number d_{ij}. The distance function serves as a model of mental arithmetic. Metric distances between the projections of stimulus points on a set of mutually orthogonal axes in a m-dimensionally organized space rest upon the parameter $r \in$ IR that determines the particular metric (e.g. Micko & Fischer, 1970). This class of metrics involves three different rules of combination, each of which predicts the rank order of similarities along mutually independent sensory attributes: For r = 1, commonly called the "city-block distance", perceived differences along the various dimensions of stimuli are merely summed or averaged to the overall similarity among the stimuli, however positive or negative.[2] For r = 2, called the "Euclidean distance", subjective differences are squared and summed, which equals the sum of similarities among the stimuli differing with respect to perceptually distinct, uncorrelated attributes, which are to be represented by a set of mutually orthogonal axes. Once two objects in a 2D metric space differ by units x_i on one attribute and by units x_j on another, their overall subjective difference or overall similarity is given by the Pythagorean theorem, $(x_{ik}^2 + x_{jk}^2)^{1/2}$. For r = ∞, i.e., if r approaches infinity, then we get what is called "supremum

[2] Scaling methods focus on the perceptual organization of the stimuli of interest, and the search for fundamental scales underlying perceptual judgments in psychology has been identified with the search for an additive operation (Luce, Bush, & Galanter, 1963, 16).

distance". Its proximity for every pair of stimuli is merely determined through its largest perceptual difference, namely $d_{ij} = \max |x_{ik} - x_{jk}|$. The kinds of spatial models described above have been found to have predictive power. Writers like Householder and Landahl (1945), Landahl (1945), Attneave (1950), Messick (1956), Torgerson (1958), Restle (1959), Shepard (1958), Garner (1962), Hyman and Well (1968) advanced the idea that if the stimuli are analyzable, i.e. obvious and natural (e.g. geometrical figures), then the city-block metric should be the best model to explain dissimilarity judgments. If the stimuli are integral, i.e. homogeneous (e.g. color patches), then the Euclidean combining rule should be more appropriate. If dimensions are integral, they do not reflect the immediate perceptual experience of the subject (Garner, 1974). By knowing the dimensions of stimuli, Attneave (1950) favored an additive model underlying the perceptual or cognitive structure of a set of stimuli. By separately judging sets of stimuli varying along subjectively independent sensory attributes, an additive model can provide a better fit to the data than the Euclidean model. Similarly, Shepard (1964) made a distinction between similarity as a basic, perceptual relation between unitary stimuli with what Torgerson (1965) calls a multivariate attribute, for which the Euclidean model is required, otherwise the similarity as a derivative, cognitive relation between analyzable stimuli into perceptually distinct components, which is viewed as a qualitative, class variable, and so city-block model is suitable .[3]

The development of MDS pioneered by Shepard (1957, 1962) and Kruskal (1964a,b), and later extended by a number of researchers with a range of sophisticated measurement techniques to apply to data obtained from both preference ratings and similarity judgments. Using the "unfolding" model (Coombs, 1950) or "points-of-view" model (Tucker & Messick, 1963) or "three-mode factor" model (Tucker, 1972) several studies (e.g. Bennett & Hays, 1960; Cliff, 1969; Carroll & Chang, 1970; Green & Carmone, 1969), beyond the scope of this paper, showed that the multidimensional solution generally resulted in two or three dimensions shared by all stimuli tested, and that interindividual differences, i.e. preferential choice data in terms of Coombs classification (Coombs, 1960, 154), may be related to the same dimensions, called the "common stimulus space". This idea for explaining individual differences was first introduced by Horan (1969). Individual differences models (for a more details see Cliff, 1968) look for interindividual differences that emerge in the weights applied to the dimensions and in the angles between dimensions. All subjects use the same set of dimensions but differ with regard to the relative importance or weight of these dimensions in influencing the judgments of similarity, thereby the directions representing subject's projections of stimuli on a vector correspond to individual's points-of-views with regard to the coordinates in a common stimulus space. Analysis

[3] Along with Hyman and Well (1968), the analyzability of the stimuli is understood as the phenomenological obviousness or perceptual distinctness or separability of the component dimensions.

of subject's preferential data called I-scales yields a joint space called J-scale with individual's vector directions and points of stimuli. In a weighted MDSCAL, i.e. INDSCAL, an algorithm advanced by Carroll and Chang (1970) for the purpose of carrying out three-way MDS, performs a metric MDS. INDSCAL can be expressed in the following general form:

$$d_{ijn}^{(t)} = \left[\sum_{k=1}^{m} w_{nk} |x_{ik} - x_{jk}|^r \right]^{1/r} \tag{1.2}$$

where $d_{ijn}^{(t)}$ is the distance between stimulus i and stimulus j for subject t, w_{nk} is the weight of dimension k associated with source N ($w_{nk} \geq 0$), and x_{ik} and x_{jk} are the coordinates of two stimuli on dimension k. The square of a subject's weight on a dimension indicates the proportion of variance of his similarity data that can be accounted for by that dimension. In contrast to MDSCAL, the dimensions obtained directly from INDSCAL are usually interpretable without rotation (e.g. Carroll & Wish, 1974a). INDSCAL reveals a group space that reflects the perceptual representation of stimuli as well as the relative salience for each source or subject (e.g. Howard & Silverman, 1976). INDSCAL assumes that overall similarity is a decreasing linear function of metric distance in any latent psychological space. Additionally, there are a number of dimensions pertaining to the similarity judgments, which should be common to all individuals (e.g. Wish et al., 1970; Carroll & Wish, 1974b; Miller & Carterette, 1975; O'Hare, 1976; Milord, 1978). As a special case of extended INDSCAL model with common and specific dimensions, Winsberg and Carroll (1989) as well Winsberg and De Soete (1993) proposed a weighted INDSCAL model, CLASCAL, in a general form:

$$d_{ijn}^{(t)} = \left[\sum_{k=1}^{m} w_{nk} |x_{ik} - x_{jk}|^r + v_n (s_i + s_j) \right]^{1/r} \tag{1.3}$$

where s_i is the square of the coordinate of stimulus i along the dimension specific to that stimulus, and v_n is the weight given by source N to the whole set of specific dimensions ($v_n \geq 0$). The specific dimension, s, refers to the square of the perceptual strength of such a dimension (e.g. for a general review see McAdams et al., 1995).

2.1.2 Metric Conditions

If dissimilarity judgments can be mapped by an ordinal transformation onto distances of Minkowski spaces, some qualitative conditions must be satisfied (e.g. Tversky, 1977; Schönemann & Borg, 1983). Young and Householder (1938) established three metric properties to assure that ordinal similarity data could be embedded in metric distances between pairs of stimuli in a dimensionally organized space. There are three equalities in their concept of metric spaces: A metric space is a pair (K,δ) where K is a set of points $x, y \in K$ and $\delta(x,y)$ is a real nonnegative function on K % K and satisfying the conditions for a distance function (e.g. Coombs, 1960, 158):

(i) Minimality: $\delta(x,y) = 0$ iff $x = y$
(ii) Symmetry: $\delta(x,y) = \delta(y,x)$
(iii) Triangle Inequality: $\delta(x,z) \leq \delta(x,y) + \delta(y,z)$

Such equalities fix a metric scale in that they are necessary but not sufficient for its specification. Minimality, sometimes called nonnegativity, holds if and only if (iff for short) that the ordinal measures δ are equal on diagonals of proximity matrix. The distance of any point to itself is always zero, and any distance between two different points is greater than zero. Symmetry requires that the distance does not depend on the order of the points. If only minimality and symmetry of metrics such as (i) and (ii) except triangle inequalitiy axiom such as (iii) are to be fulfilled, then it would be semimetric (Shepard, 1974). The triangle inequality is violated if the dissimilarity judgments are based on different judgment criteria. If x, y, and z differ only along one dimension, dissimilarity ratings can be subadditive in the sense that $\delta(x,y) + \delta(y,z) > \delta(x,z)$ (e.g. Schönemann, 1994), which is highly correlated with violations of minimality. Subadditivity of dissimilarities could be explained as largely due to the fact that subjects tend to stay away from the lower response bound (i.e. 'identical') of the rating scale, thus in effect adding a positive constant to all distance estimates δ (Staufenbiel & Borg, 1987).
If distance estimates do not satisfy minimality and symmetry, they cannot be represented by Minkowski distances (e.g. Schönemann & Borg, 1983). Besides minimality $0 = d(x,x) = d(y,y) \leq d(x,y)$ and symmetry $d(x,y) = d(y,x)$, a further qualitative requirement, ultrametric inequality $d(x,z) \leq \max[d(x,y), d(y,z)]$, evolved by Johnson (1967, 245), must be fulfilled for the data to hold (e.g. Fahrmeir & Hamerle, 1984). If any three points x, y, and z lie collinearly on a straight line, such as a dimension, in psychological space, then the segmental additivity axiom $d(x,y) + d(y,z) = d(x,z)$ is fulfilled (Gati & Tversky, 1982). When the metric axioms as such are supported by the data, the application of metric models to the data might be justified (Beals et al., 1968, 128). Symmetry would not hold if one object is a prototype and the other one a variant of this prototype. When only minimality and symmetry, but not the triangle inequality are to be satisfied, it leads to semi-metrics. If

asymmetries arise in dissimilarity ratings, it implies that a subject possesses no reliable basis for stimuli being judged. In other words, the subject is unable to make a weak order relation in that connexity and transitivity are thought of as requisite and sufficient conditions for the empirical system.[4] Ultrametric inequality refers to the fact that the distance between the pairs of stimuli is not larger than its distance to a third object at all. In a sense, this type of scale has to do with the ordered metric (Coombs, 1950) in which stimuli constitute an ordinal scale, but distances between them are partially ordered.

2.2 Nonmetric Multidimensional Scaling

The conversion of responses on sets of stimuli on an ordinal scale into numbers that are elements of an interval or ratio scale onto which point representations of these stimuli are to be located is a central issue in psychometrics. Along with Green and Carmone (1969, 1972), metric MDS is based on the interval as well as the ratio scale, whilst nonmetric MDS is subjected to the ordinal scaled data. As in Torgerson's classical, metric MDS (Torgerson, 1952, 1958), similarities are assumed to be related to metric distances (distance model) among stimuli in a dimensionally organized space according to some distance functions (space model). Very briefly, its main features may be outlined as follows: The first step is to construct sets of similarities and to calculate their metric distances. The second step is to attain the global minimum departure from the monotonicity to circumvent local Stress minima in such a manner that metric distances are reduceable to a set of mutually orthogonal axes on as few dimensions as possible. The principal axes are the scales wanted, and they are then interpreted as psychologically meaningful dimensions or features that underlie the perceptual judgments. Ordinal MDS makes no a priori assumptions about the physical bases of the dimensions that underlie the perceptual representation. The details on how to interpret the sensory attributes revealed in a geometric space dimensionally and/or nondimensionally can be found in a contribution by Gigerenzer (1977). The orthogonality of principal axes assumes the independence of sensory attributes, which, of course, is merely a special case for a psychological space that is free of direction. In this connection, the relations between the spatial models and the differential attention reflecting the relevance of sensory attributes have been studied by Wender (1969) and by Micko and Fischer (1970).

[4] A two-digit relation is called connex, when, for every object pair, either a b or b a can be produced, and transitive iff when a b and b c then a c (see e.g. Tversky & Krantz, 1970). Referring to the intransitivity axiom in perceived similarity of sounds, to the clarity of the matter, sound S_1 and sound S_3 are not similar to each other iff when sound S_1 and sound S_2 are similar in respect to attribute A_1, and to nothing else, sound S_2 and sound S_3 are similar in respect to attribute A_2, as well as to nothing else.

Nonmetric MDS advanced by Kruskal (1964a,b) calls merely for a monotonic relationship between the ordering of similarities and the rank order of metric distances in any dimensionally organized metric space. The characterization "nonmetric" refers to the fact that Kruskal's procedure is capable of constructing a configuration of points in a metric space using no more than the ordinal properties of the data in a dissimilarities matrix. With a single, complete set of N(N − 1)/2 proximity measures of pairwise similarity ratings among N objects in interest, N % N similarity matrix in which N rows and N columns correspond to the same N objects. The entries in this matrix may be individual data or group data. The proximity scores are averaged and transformed to a scale ranging from 0 to 1. Metric distance is expressed as a function of ranked dissimilarity that is the starting point of Kruskal's technique. In Kruskal's procedure, N points are positioned in a m-dimensional space in such a way as to achieve the best possible approximation to a monotonic relationship between N(N − 1)/2 metric distances and an empirically obtained ordering of pairs of stimuli with respect to similarity or to dissimilarity. Such a space thus obtained has the status of a unique property of the recovered information about stimulus points. Suppose dissimilarities behave like psychological distances, some latent psychological space follows necessarily. If, post hoc, distance geometries predict all that can be said of the ordering of pairs of stimuli with respect to dissimilarity, then the observed dissimilarity ordering of the pairs is assumed to reflect the order of their metric distances at the level of analysis involved (Shepard, 1966). The decision as to the measure of similarities or dissimilarities rests upon the purpose for which it was generated (Bäumer, 2000).[5] As for the similarity measure on the basis of interval and/or ratio scale, Pearson's correlation coefficient is favoured, especially when similarity should be stipulated as an extent of linear coherence between two variables. As far as the rank order of proxmities is considered, the dissimilarity measure is indicated.

2.2.1 Minimalization Problem

Elaborated on the Shepard program (Shepard, 1962), a somewhat ad hoc iterative method to minimize Stress for Minkowski distances was pursued by Kruskal (1964a) (e.g. for problems of local minima see Shepard, 1974). Very briefly, the rationale of overall minimum solution may be outlined as follows: One tries a number of random starting configurations with zero Stress and takes the value of Stress that repeatedly turns up with the same smallest value departure from monotony. This delivers further configurations with which the fitting process continues. To get a measure of departure from good fit in

[5] Discussions on the issue related to various similarity measures can be found in the papers by Bray and Curtis (1957) and by Goodall (1966).

spaces of decreasing dimensionality, the fitting will be continued insofar as Stress falls short of a very small given number (Cox & Cox, 1990, 1992). One refers to a solution from which no small change is an improvement and yet which is not the optimum solution (Kruskal, 1964b, 118). The process of iterations comes to an end, corresponding to the case where the difference between the coordinates of iteration i and i – 1 would be smaller than a local minimum criterion generally set to 0.005. Shepard (1962, 1974) took at least 20 random trials of any initial configuration as necessary to land at an overall Stress minimum. Lingoes and Roskam (1973) showed that Kruskal's method is much more vulnerable to local Stress minima than Guttman's method.

2.2.2 Loss Function

A series of questions, such as the appropriateness of a particular mapping of the data into distances, or the true metric and dimensionality of the psychological space, rely exclusively on the global loss, Stress. Kruskal presented a procedure for minimizing a loss function such as Stress with which to test the nonmetric hypothesis allowing for a monotonous response function that relates the computed metric distances to dissimilarity ratings. The degree of random fluctuation to which an ordering of metric distances result from some unknown monotone distortion of an ordering of dissimilarities is expressed by the root-mean-square residual:[6]

$$Stress_{formula-1} = \sqrt{\frac{\sum [d_{ij} - \hat{d}_{ij}]^2}{\sum d_{ij}^2}} \qquad \text{(Kruskal, 1964a)}^{7} \qquad (1.4)$$

The numerator by Kruskal's Formula One, quite similar to the residual sum of squares, is meant to show the degree of monotone distortion of observed dissimilarities from true dissimilarities (d-hat), while the denominator makes a normalizing factor for the invariance of fitting degree with regard to expanding or diminishing the configuration. Kruskal's analysis repeats itself until the final configuration with some metric and dimensionality would be accepted as best, having the smallest Stress value, hence the fitting is the greatest. This procedure decreases the final Stress values slightly. The optimal dimensionality would be estimated by inspecting the addition of dimensions

[6] In analogy, Stress in MDS can be considered to be the standard error in regression analysis.

[7] Kruskal's Formula One reveals a more sensible configuration with fewer iterations. For reasons of disposing a possible degenerate solution, exceptionally, Kruskal's Formula Two,

$$Stress_{formula-2} = \sqrt{\frac{\sum [d_{ij} - \hat{d}_{ij}]^2}{\sum [d_{ij} - \bar{d}]^2}} \qquad (1.5)$$

has been used in practice, though at the cost of higher Stress values.

which could lead to the reduction of Stress marginal. The range of Stress lies between 0 and 1, and the closer a value of Stress comes to 0 the better the configuration. Table I-1 indicates the basis for the recommendations of Stress for the final configuration, which will be called the local Stress minimum criterion.[8] Very low Stress values suggest that the stimulus spacing is very coarse.

Table I-1: Evaluation of Stress

Stress values	**goodness-of-fit**
Stress > 0.3	any positioning of points
0.2 < Stress ≤ 0.3	deficient (for N < 50)
0.1 < Stress ≤ 0.2	unsatisfactory
0.05 < Stress ≤ 0.1	satisfactory
0.01 < Stress ≤ 0.05	good
0.00 < Stress ≤ 0.01	very good
Stress = 0.00	perfect

There are no standards for evaluating such a global loss value, for the distribution of Stress is unknown. However, special Monte Carlo methods show that the random function of Stress is specified as a function of both the number of stimuli N and the number of dimensions t (Stenson & Knoll, 1969; Klahr, 1969; Wagenaar & Padmos, 1971; Sherman, 1972; Spence, 1972; Spence & Graef, 1974; Graef & Spence, 1979). Main features of Monte Carlo algorithm [9] evolved by Wagenaar and Padmos (1971) may be outlined as follows: The first step is to produce a random configuration of N points (N = 8, 10, and 12) in t dimensions (t = 1, 2, and 3), by a random sampling of (N × t)-dimensional configuration space from a uniform distribution on the open interval 0-1000. The second step is to construct sets of synthetic dissimilarities retained from t-dimensional model space. The third step is to calculate error-perturbed metric distances according to a model using the Euclidean metric (e.g. Graef & Spence, 1979):

[8] Kruskal's evaluative labels of Stress (Kruskal, 1964a, 3), i.e. 0.2 = poor, 0.1 = fair, 0.05 = good, 0.025 = excellent, 0 = perfect, is, due to its vagueness, of no use in practice. The evaluation of Stress depends on the number of stimuli involved, frequently Stress rises with the number of objects taken into consideration.

[9] The term "Monte Carlo" has been coined by Ulam. It is not central to the concerns of this paper that the reader understands how such Monte Carlo designs are constructed and used in the multidimensional context. For a general review of Monte Carlo methods applied to various issues in multidimensional scaling, see Spence (1983).

$$d_{ij}^{\ e} = \left[\sum_{k=1}^{2} \left(x_{ik} - x_{jk} \right)^2 \right]^{1/2} \cdot N(1, \sigma_x^{\ 2}) \tag{1.6}$$

Each distance is multiplied by a random normal deviate with a mean of one and a standard deviation of σ_x.[10] The input data are then analyzed in one through five dimensions. The Stress thus retained is compared with that of the (N + 1)-dimensional configuration, and so on. Once the (N + 1)-dimensional solution has a better fitting, say no monotonous distortion, then (N + 1)-dimensionality can be taken as the appropriate dimensionality for any number of stimuli of interest. This Monte Carlo procedure is replicated a number of times so sufficiently as to get stable estimates of variance, i.e. until the "predicted" Stress does not deviate from the obtained value. On the basis of analysis of synthetic dissimilarity matrices from random configurations consisting of from 12 to 48 points in one through five recovered dimensions, Spence & Ogilvie (1973) could construct a table of expected Stress values. Wagenaar and Padmos (1971), in an 8 × 8 dissimilarity matrix, proposed a Stress of 0.4 for one dimension, 0.34 for two dimensions, and 0.2 for three dimensions. Non-significance of Stress caused by error rests on the number of dimensions, i.e. the more the dimensions extracted, the lower the error should be, therefore the Stress will be lowered. With 9 points in 3 dimensions, Klahr (1969) proposed a Stress value less than 0.075 as a lower bound on the number of stimuli.

3. Experimental Studies

We may now proceed to report two experiments of our own. In each, a group of vehicle interior noise was chosen. These were presented in pairs to the subjects, Ss for short, who rated each pair on scales of preference and similarity. A paired comparison evaluation technique (e.g. David, 1988) used to determine the preference for sounds is currently the most favorite sensory method for sound quality engineering. Two major aspects of perceptual ratings will be discussed. First, we shall be interested in mapping sensory spaces with regard to pairwise preference ratings and pairwise similarity judgments. A technical procedure for determining the "true" dimensionality of geometric representations of two empirically obtained dissimilarity

[10] Thurstone speaks of what he calls 'discriminal process' by which the internal effect of stimulus can be converted into a random variable wiht its range being the real numbers. If a normal distribution of a random variable is assumed, its mean and variance may be considered parameters to be estimated. The Euclidean model multiplied by N(0,1), which is a normally distributed random variable with a mean of zero and a unit variance analogous to what is called multidimensional Thurstone's Case V (Torgerson, 1958; Luce, Bush, & Galanter, 1963).

matrices being analyzed is presented. Secondly, to be checked is the hypothesis assuming that pairwise preference ratings can be a subset of the perceptual structure brought about by pairwise similarity judgments. Experiment I evaluated the vehicle interior noises in terms of liking and disliking, and Experiment II dealt with perceived similarity. The results were analyzed dimensionally using the nonmetric scaling procedure and nondimensionally using the additive tree procedure. A 2D configuration, well matching the nondimensional clustering results, was found to be suitable. In comparison with the similarity structure, the perceptual dimensions retained in the preference structure were found to be reflected in those pertaining to the similarity judgments.

3.1 Experiment I : Preference Scaling [11]

With the goal of capturing the judgment criteria by which to make preference ratings on different vehicle interior noises, we conducted a psychological scaling experiment on the perception of vehicle noises. The paired comparison method worked out by Thurstone (1927), which was used for preference ratings in the present study, or methods of successive intervals (e.g. Edwards, 1952; Messick, 1956; Diederich et al., 1957), are models for collecting data (e.g. see Coombs, 1960). The axiomatic foundation of the preference relation can be found in papers by Coombs (1950) and by Luce et al. (1963).[12] With N stimuli, a total of $N^2 - N$ ordered pairs can be formed (Tucker & Messick, 1963). To be answered are those questions as to whether the preference ratings meet the condition of symmetry for the applicability of nonmetric scaling analysis, whether the preferences of pair relations of noises are decomposable into differentiable psychological dimensions, how few sensory attributes are considered necessary for the preference assessment.

3.1.1 Method

To be examined is whether the dissimilarity ordering of pairs of sound sources in the preference data is to be matched with the rank order of their

[11] Some of these results were reported at the First KRISS International Autumn Session in Ergonomics, September 2000, Taejon (Choe, 2000a).

[12] If A_1 is a set of subjects, $a \in A1$, A_2 a set of stimuli, $\alpha, \beta \in A_2$, and T is defined as the ternary empirical relation $T(a,\alpha,\beta)$ iff a likes α at least as much as β. Let A be the union of A_1 and A_2, then the Coombs system $<A,A_1,A_2,T>$ is called an empirical preferential system. In a numerical preferential system $<N,N_1,N_2,S>$, N, N_1, N_2 are sets of real numbers, $x \in N_1$, $\zeta, \omega \in N_2$, $N = N_1 \cup N_2$, and S is a ternary numerical relation such that $S(x,\zeta,\omega)$ iff $|x-\zeta| \leq |x-\omega|$. When N_r is a set of r-dimensional vectors $\mathbf{x} = (x_1,x_2,...,x_r)$, where each x_i is a real number and $S_1,...,S_N$ are numerial relations on the vectors in N_r. Then the empirical preferential system $<A,A_1,A_2,T>$ is an r-dimensional vector system (Luce, Bush, & Galanter, 1963, 45-47).

metric distances recovered in any dimensionally organized space imposed by Kruskal's analysis, which Minkowski power metric arrives at an optimal rule of cognitive combination that reflects the rank order of preference data at best, and what the organization pattern in the data looks like. Finally, it will be discussed how to interpret the multidimensional solution of principal axes in the final configuratoin adequately.

3.1.1.1 Subjects

A total of thirty subjects with an equal number of male and female subjects, aged 27 years on average, pooled over different faculties at the University of Oldenburg, participated in the preference ratings. All Ss stated having normal hearing on the questionnaire. Sound preference judgment tasks are familar to most Ss, hence, even if Ss have not heard particular car noises, they will most likely have some opinion regarding it.

3.1.1.2 Materials

Using physically simple stimuli with familiar characteristics we may indeed be able to determine the characteristics of the data that are essential for successful construction of the configuration, and how they combine with each other. Here, we are mostly interested in complex sound sources with which we are more or less faced in our daily life. Once selected sound samples, which are appropriate in cognitive studies, are reasonably representative of the class, then the psychological distance measures between the stimuli retained from nonmetric multidimensional scaling would reveal the fundamental psychological dimensions of stimuli underlying the class. The corpus of heterogeneous stimulus materials related to a single class of sound samples contains 15 representative sequences of commericially available vehicle interior noises gleaned from pickups of Mimic Sound Expression for Vehicle Noise of the Society of Automotive Engineers of Japan Inc. (1994).

Table I-2: Classification of materials
(mimic sound expression for vehicle noise, 1994)

Track No. (CD No.)	**Classification (Symbol)[max. dB]**	**Phenomenal Description**
01 (A 30)	Generator engine noise (GG)[86 dB]	This whine noise is conspicuous when racing is generated by cooling fan of generator. Frequency depends on number of blades.

continued from Table I-2

Track No. (CD No.)	Classification (Symbol)[max. dB]	Phenomenal Description
02 (A 32)	Cooling fan engine noise (AVG)[90.7 dB]	This boom noise is generated when engine cooling fan turns at high speed. The higher the speed is, the louder the noise becomes. And this noise is based on the frequencies of orders of blades.
03 (A 34)	Turbo engine combustion noise (VMG)[86.9 dB]	This chirping noise is a mechanical noise generated from injection pump (as recorded with engine idling).
04 (A 42)	Injection pump engine noise (ASG)[86.59 dB]	This rattling noise is generated from the transmission case of a manual transmission vehicle standing in neutral with clutch engaged. Speed irregulartiy of the flywheel are transmitted by the transmission and cause driving and driven gears to have their teeth hit each other. Resultant vibration is propagated to the transmission case to produce the rattling noise.
05 (A 8)	Booming engine noise (DG)[91.7 dB]	This booming noise is typical on vehicles with automatic transmission. It occurs with rapid growth of sound pressure because of high engine load, from standing start or during acceleration from low rpm. It sounds 'moan' when the sound pressure level of the 1st, 2nd and 3rd engine firing order are high and almost equal.
06 (A 62)	Noise from transmission case (TGG)[91.7 dB]	Thie whine noise occurs at all types of transmission. High frequency gear meshing noise occurs from transmission gear, differential gear while cruising. Vibration occured by gear meshing error, is transmitted to transmission gear case and differential gear case and then the noise is generated.
07 (A 88)	Clutch operating transmission noise (KG)[89.9 dB]	Huming noise is heard from fire wall by clutch pedal operation. Noise may be generated by the vibration through the clutch system from engine.
08 (B 28)	Brake squeak transmission noise (BQG)[82 dB]	This booming noise is a pulsating noise generated in the air intake and exhaust lines when the exhaust brake valve is closed, and is mainly emitted as a transmitting noise from the suction port of the air intake duct and heard in the cab as a solid borne noise caused by vibration in the intake duct.
09 (B 54)	Road noise (W1G)[90.09 dB]	At high speed, Karman vortexes in the rear of rod generate wind noise.Wind noise frequency is related to vehicle speed and diameter of antenna. Vibration of antenna and wind disturbances change the frequency.
10 (B 56)	Road noise (W2G)[88.2 dB]	This wind leak noise is the high frequency wind noise generated by air flow through a door weather strip opening, when driving against strong wind.
11 (B 62)	Tyre noise (R1G)[89.09 dB]	While crusing on rough roads, suspension system is vibrated by the tire response to road irregularities, and the suspension system vibration excites body panel resonance vibrations. As a result, road noise is generated by the panel vibrations in the cabin. You can feel very little vibration.
12 (B 64)	Tyre noise (R2G)[86.7 dB]	While crusing on rough roads, suspension system is vibrated by the tire response to road irregularities, and the suspension system vibration excites body panel resonance vibrations. As a result, road noise is generated by the panel vibrations in the cabin. If an irregularity is large, it feels like fluctuating pressure at ears.

continued from Table I-2

Track No. (CD No.)	Classification (Symbol)[max. dB]	Phenomenal Description
13 (B 74)	Air-conditioner compressor vibration noise (KAG)[91.3 dB]	Air conditioner (A/C) compressor's operation causes compressor's vibration because of pressure pulse of A/C gas. The vibration is transmitted to the vehicle body and makes A/C operation noise.
14 (B 78)	Instrument panel radiation vibration noise (IBG)[92 dB]	The instrument panel radiation noise is generated by vibration of engine and transmission transmitted to the instrument panel through speed meter cable.
15 (B 94)	Gear shift lever vibration noise (GHG)[84.59 dB]	This rumbling noise is mechanical noise generated by vibration of the gear shift level and the floor panel, where the level is fixed, excited by vibration of the transmission control linkages at a gearshifting operation.

Binaural recorded sounds inside and outside of motor vehicles referred to the ISO-Recommendation (ISO/R 362, ISO 5128), in which the measurement conditions were pinned down to make the results reproductible. For calibration purposes, each recording of a sound was preceded by a recording of a 1 kHz pure tone at a level of 60 dB near the microphone.[13] The choice of stimulus materials was for two reasons: (1) The use of synthetic or synthetically modified real engine noise through the use of filter functions can lead to false assessments about the meaning of objective parameters, which are shown to be an useable variant explaination of perceptual ratings for real engine noise (e.g. Beidl & Stücklschwaiger, 1997). (2) The 15 sound sources, commonly encountered in cars, give a fair representation of vehicle interior noises (e.g. Brunswik & Kamiya, 1953).[14] In a sense that sensory attributes that make a sound typical for a class are such which make members of one class distinguishable from members of a different class, a set of sound sources chosen for this investigation can be partitioned into five noise classes, as is listed in Table I-2. The vehicle noises were edited to minimize the temporal influence on perceptual rating, and thus the duration of sound samples was equalized at mean time of 15 seconds.

3.1.1.3 Apparatus

For the purpose of psychological scaling of acoustic preferences a special computer program NOise eVALuation, or NOVAL for short, was developed, which runs on AT/AT-compatible PCs. The essential software properties are that the program was configurated with the aid of BORLAND DELPHI 4.0

[13] Aurally adequate recording and analysis techniques using the Artificial Head Measuring System have been used since 1980 for the purpose of assessing acoustic quality and designing sound (Genuit, 1997).

[14] Although there is no sensible procedure for choosing 'typical' samples in this sort of hearing situation, it would be desirable to investigate perceptual judgments of natural objects with which subjects are initially familiar. As was already pointed out by Brunswik and Kamiya (1953), results with "ecological validity" may be obtained only by the use of stimulus materials which are representative of the real situations to which one wishes to generalize.

(Standard Version), INPRISE Corporation (1998), and it runs under Windows NT Version 4.0.

Table I-3: Arrangement of delphi components
(visual component libraray, VCL)

Classes	Tasks	Components
Standard	Text Contol Elements	Edit, Memo, MaskEdit, RichEdit
Win32	Special Input Control Elements	ScrollBar, TrackBar, UpDown, Hot-Key, Splitter
Additional	Interfaces	Button, BitmapButton, SpeedButton, CheckBox, RadioButton, ToolBar, CoolBar
Standard, Win32	Handling of Lists	ListBox, ComboBox, TreeView, ListView, ImageList, CheckListBox, DateTimerPicker
Standard, Win32	Grouping of Compo-nents	GroupBox, Panel, PageControl, Ra-dioGroup, ScrollBox, HeaderControl
Standard, Win32	Visual Feedback	Label, ProgressBar, StatusBar
Data Control	Tabular Notices	DrawGrid, StringGrid, DBGrid
System, Win32	Graphical Notices	Image, Shape, Bevel, PaintBox, Ani-mate
Dialog	Standard Dialog Field of Windows	Data Operations

The use interface is 795 % 575 pixel, a display centered in the monitor with a 1024 x 768 screen. DELPHI is an object-oriented programming language for rapid application development. The standard programming language underlying DELPHI is OBJECT PASCAL. The development of DELPHI was recently updated as DELPHI 5.0 which is available on the market. Thanks to the integrated development environment, visual tools (Visual Component Library, VCL), as is listed in Table I-3, can be tailored to user's different purposes of applications. The compiled executive file NOVAL.EXE, which handles user interface, is working regardless of Dynamic Link Libraries (DLL). NOVAL interface connects the computer with CD-ROM, thus all stimulus signals stored on a Compact Disk (CD) can be fed analogously from CD-ROM via Sennheiser headphone HD 540 reference II in a hemi-anechoic chamber, by order of sound presentation through pressing the button. Maximum decibel of sound sources based on CD-ROM is bound up to about 96 dB (see dB Max values in Table IV-1). The converter of CD-ROM scans 65 536 steps ($\approx 2^{16}$), so 16 bit-resolution (187 KB/sec) renders to amplify the stimulus signals. Sounds were played back by headphone, the controller being

held constant during presentation of sound sources, so that the equalization problem relating the aurally adequate sound evaluation is settled.
To the detriment of headphone reproduction for sounds, the bodily feelings or realistic feelings of presence by low frequencies are missing in binaural representation in headphones. However, Khan et al. (1998) showed that there was no significant difference in the judgments of annoyance between headphone and loudspeaker reproduction for 10 diesel engine sounds. As with the program module for preference scaling (see Fig. I-1), a bipolar rating scale with interval character is implemented in the program, as well as verbal labels for relatively rough adjustment at the beginning of each experimental session.

Fig. I-1: Program interface for preference evaluation

Only 1 step increments or decrements are allowed. This gives a 100 step rating scale. Ss entered the judgment by moving a track bar on the program interface that represented a continuous bipolar rating scale ranging from unequivocal for the first sound (coded –50), a bit more for the first sound (coded –25), equal for both sounds (coded 0), a bit more for the second sound (coded +25), to unequivocal for the second sound (coded +50). When the perceived value between two sound samples being compared increases and is positive, then it can be stated that the second sound is preferred over the first sound. The same holds for the negative case. The Ss were asked to make preference ratings on 40 possible stimuli combinations of 15 vehicle interior

noises. A total of 210 pairs of 15 sound stimuli in preference ratings were collected. It may not be desirable to collect complete pair comparisons from a subject due to the fact that 210 pairs to be evaluated in a single test can absolutely and uniquely be too onerous a burden for a subject. In order to enhance the adjustment precision each step has ten equidistant undervalues. Such a scaling technique is somewhat like Category Partitioning Scale developed by Heller (1985) and Hellbrück (1986, 1991). In a sense, this rating scale may be thought of as a hybrid between the magnitude estimation and the direct scaling.

3.1.1.4 Procedure

The like-dislike ratings of vehicle noises were conducted in a slight sound-absorbing laboratory at the MUB via earphones. The position effects can be controlled by means of random generation of sequence of sound samples. The preference experiment is analogous to the method of absolute judgement in psychophysics, which is regarded as a modification of the method of standard stimuli (Weber & Zener, 1928). It amounts to omitting the standard stimulus, so that a single stimulus is presented on each trial. As the S becomes acquainted with the range of stimuli involved, his/her responses settle down to asymptotic levels. It is as though the S defined his/her own standard stimulus for the given set of comparisons and that he/she is able to focus this image well during the course of experiment. The course of the experiment can be illustrated in Fig. I-2. Prior to conducting the main session, every S had a chance to be familiar with the instruction of experiment, and to clarify the vagueness in technical matters on the part of program applied. To achieve this, three practice trials chosen randomly from the set of pairs were run. For each S a different random order of presentation of the pairs of sound samples was brought into work. The main part of experiment took approximately 45 minutes. Each sound sample was presented for 15 sec. All listeners were asked to hear the vehicle noises in every pair in terms of the question “which of two sounds do you prefer?”, and to indicate the strength of their likings or dislikings on a judgment scale provided. It was stressed that the task merely relies on overall impressions that pop into the minds of Ss. This would allow a listener’s bias to come into play. Each sound pair could be replayed at will until the S was satisfied with his/her judgment. Every time the judgment was entered, the next trial was automatically prepared. Once the experiment was completed, the Ss were given a questionnaire in that every S was asked to name descriptive adjectives or judgment criteria which were relevant for pairwise preference ratings. The instruction of how to do the task was mounted as a button above the program interface as referred to Block I.

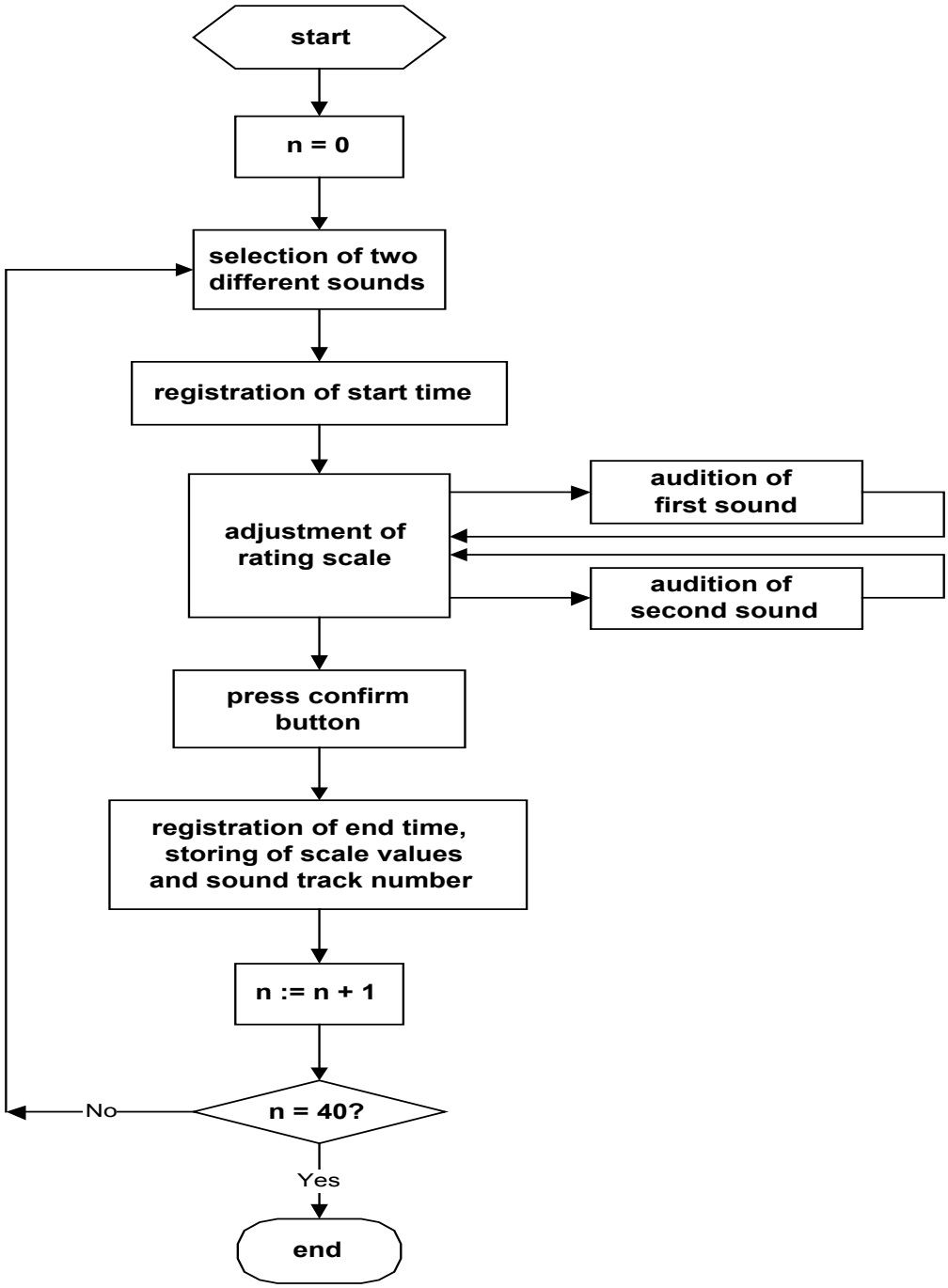

Fig. I-2: Event handling of sound assessment

Block I: Instruction for the like-dislike ratings in Experiment I

Thank you for your readiness to take part in my experiment. The experiment will take 45 minutes. In the following test you will hear two sounds from cars. Your task is to compare the two heard sounds. Please proceed as follows:

Click on the left button for the first sound with the left button of the mouse, and listen to it. After having heard the first sound, click on the right button for the second sound and listen again. You listen to every sound until it ends. Notice that you don't start the second sound before the first sound has ended. It should be noted that certain sounds in combination with other different sounds could occur. Every sound lasts 15 seconds. After listening to both sounds, you are to indicate which sound you prefer. For the sake of judging sounds, a rating scale is provided, which has a horizontal track bar. The rating scale at hand has 5 steps:

I prefer the first sound.
I prefer the first sound a bit more than the second one.
I prefer both sounds equally.
I prefer the second sound a bit more than the first one.
I prefer the second sound.

You can move the track bar on the whole scale with the mouse. To move the track bar, push it by draging the left button of the mouse, and set it to your desired position on the scale, or for adjustment you can set down the track bar somewhere between the steps, for example, between the "equal" and "a bit more" etc., while pressing on the range of scale. At the beginning of each session the track bar is automatically set in the middle. When you are not able to decide after having heard a pair of sounds, you can repeat each sound once again. Having made a decision, press the button "confirm". After that, you obtain a message "Your decision was confirmed!". Please press the buttton "OK", then the next pair of sounds is ready. The track bar is back to the middle of the scale. Just make your ratings on the basis of your general impression of which sound does seem more appealing to you, rather than on any other particular characteristic. The first two ratings you make will be just for practice. This practice series will familarize you with the sounds you are to rate. When having difficulties in dealing with the task, or if you lack claritiy about the course of experiment, you can either retrieve the instruction by pressing the button "instruction", or you may ask the experimentator at any time. Shortly after having solved the last pair of sounds, you will obtain a message at the end of the experiment. Then you may turn to the experimentator.

3.1.2 Treatment of Data and Results

The aim of nonmetric multidimensional scaling is to delineate the effective stimulus dimensions and to obtain a spatial representation of stimuli. Nonmetric analysis of proximity data results in a dimensionally organized space in that the positions of stimuli in such a metric space are adjusted in such a manner that their metric distances of an extended Euclidean spatial model with a few dimensions have a monotonous relation to dissimilarities.

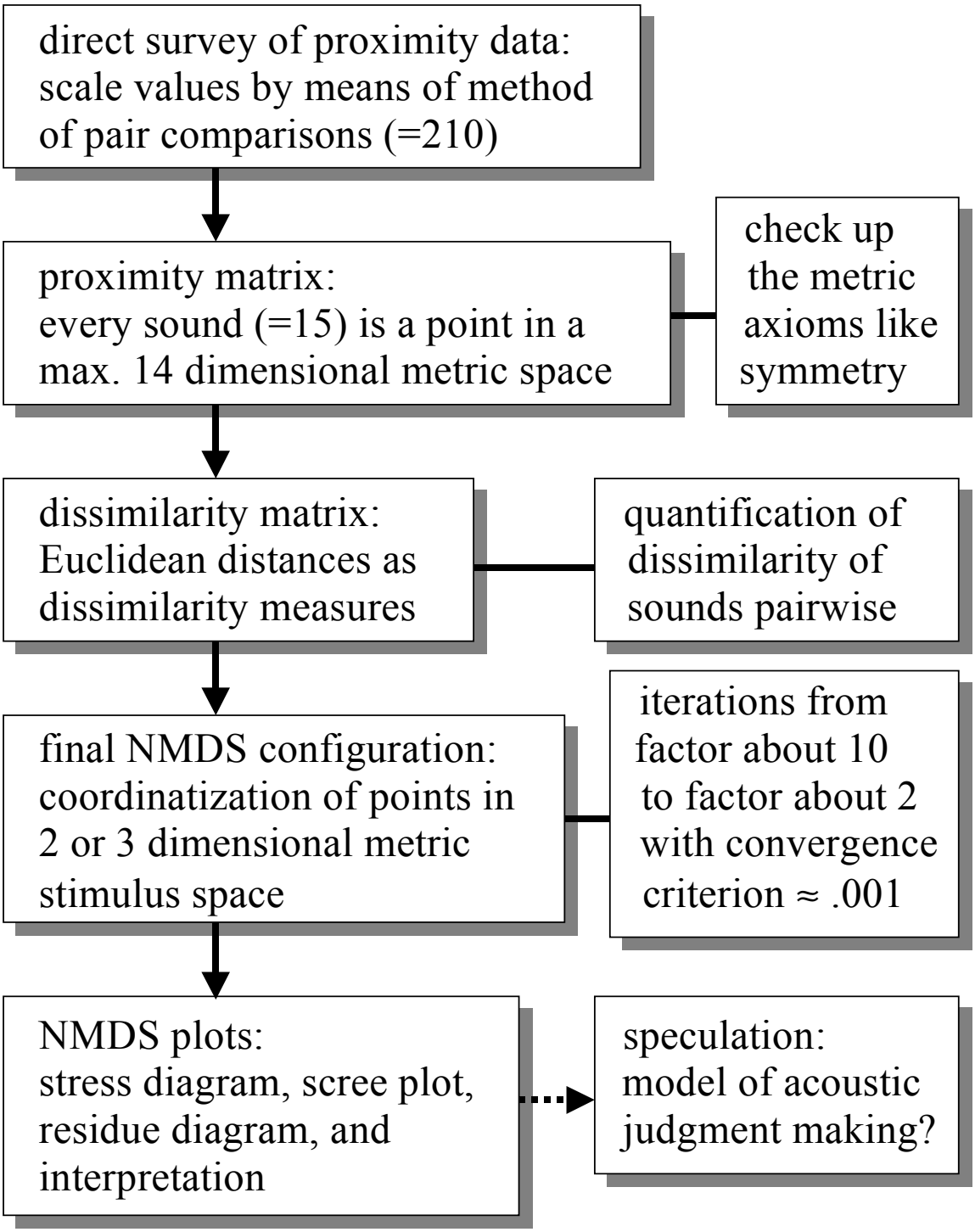

Fig. I-3: Technical procedures of NMDS

To start with, proximity values for each individual set of empirical proportions of preferences were averaged over all Ss, and this pooled proximity matrix was converted into a single set of dissimilarities which were then analyzed by a nonmetric multidimensional scaling algorithm according

to Kruskal's Formula One. The Stress of the final configuration was then evaluated. A parsimonious number of psychologically meaningful perceptual dimensions in a dimensionally organized space was selected. Such a metric analysis of ordinal data is threefold: to identify the minimal dimensionality necessary and sufficient for the data to be embedded with a minimal residual error variance, to capture mutually orthogonal axes in a psychological space, and to plot a monotonous function that relates the computed metric distances to dissimilarity ratings. The minimal dimensionality criterion for the sake of estimating any configuration of stimulus points, ordinarily all the way to (N – 1) dimensions,[15] is known for N stimulus points at the vertices of a regular simplex, with reference to the case where complete ties (equal dissimilarities) may exist.[16] The number of perceptual dimensions revealed in nonmetric scaling analysis is tantamount to the minimal number of sensory attributes for possible explanations for the unique behavior of preference data. The design of the study is straightforward and is summarized in Fig. I-3. First of all, nonmetric scaling began with gaining the proximity data. The Ss listened to every pair of sounds, p_1 and p_2, and had to judge on a rating scale ranging from –50 to +50 how they discernend it. The value of +50 implies that p_2 (second sound) was assessed to be preferred than p_1 (first sound), the value of 0 refers to the fact that both sounds were heard equally, and the value of –50 in the meaning that p_1 was preferred than p_2.

Omitting the practice series from consideration, 1200 data sets were thus obtained for pairwise preference ratings on sound samples. These were analyzed by the method of nonmetric multidimensional scaling. Having made preference evaluations from 30 Ss, each of which had to carry out 40 pairs of 15 sounds. In this manner we could secure the preference ratings for all the possible pairs in a square symmetric matrix that has a row and a column for each object and the cells containing proximities between the pairs of stimuli. Analysis of the aggregation effect through the comparison of individual solutions with an aggregate solution was not considered in our study. It is thought to be not feasible to do a separate multidimensional analysis for each subject because individuals have different multidimensional structures to be accomodated hardly (e.g. Horan, 1969). Suffice it to say here that a multidimensional solution for aggregated proximitites is usually very similar to separate perceptual structures for each individual. Using 12 adjectives drawn from the semantic differential, as a matter of fact, Anderson (1970)

[15] Kruskal's theorem is stated as saying that, for mapping proximities into both Euclidean and dominance distances, the maximum dimensionality is N – 1. For city-block distances, the dimensionality can be at most [N(N – 1)/2] – 1 (Critchley & Fichet, 1994). Lingoes (1971) describes a monotonic transformation on the proximities that reduces the dimensionality below N – 2 dimensions. Wagenaar and Padmos (1971) pointed to the fact that Kruskal's theorem holds if distances are to be represented in an Euclidean space.

[16] When two dissimilarities or metric distances build the same rank order, then one says that they are tied with each other. Ties can be interpreted as showing that there is no preference for a certain rank order of the dissimilarities (Backhaus et al., 1994). If all N(N – 1)/2 metric distances are tied, then we have a totally degenerate solution having the regular simplex in N – 1 dimensions (Shepard, 1974, 392).

found the high aggregation effect (correlation of 74%) between an aggregate solution and separate individual solutions. Later on, Anderson (1973) reported that the heterogeneous solutions based on the judgments on 10 political figures were found for nonadjective data (correlation of 56%). Similar effects have been observed using objects such as movie stars, automobiles, fruits, occupations, and ethnic groups. Regarding rectangle perception, however, Schönemann et al. (1985), Schönemann and Lazarte (1987), Lazarte and Schönemann (1991), and Schönemann (1994) reported that the aggregation of individual proximitites could not be justified because distinct strategy groups were found.

Table I-4: Scale values for paired sounds based on ratings on like-dislike continuum

$p_i \backslash p_j$	GG	AVG	VMG	ASG	DG	TGG	KG	BQG	W1G	W2G	R1G	R2G	KAG	IBG	GHG
GG		-50	5	-16,67	-35,83	-35,63	-24,5	-1,2	-7,25	-38,4	-46,67	6,2	-46,67	-25,5	-31,63
AVG	31,14		22,29	13	11,17	-14,11	21,5	30	-7,5	21,8	-18,86	-6,5	-17	-12,36	1,33
VMG	-28,57	-41,6		-17,2	-38,8	-41	-36,6	-30,33	-29,75	-25,33	-43	-19,4	-41,75	-39,6	-39
ASG	11,17	-4,71	25,5		-14,29	-31,13	-13,8	30,4	-36,4	-2	-10	18	-23,83	-40	-3,8
DG	31	-26	14,67	2		-31,25	-19	21,8	-6,88	19,67	-15	2	-12,11	-25,63	5
TGG	41,5	26,67	19	31,5	19,25		26,67	38	13,78	28	3,8	28	5	7,63	8,80
KG	16,83	-28,14	11,57	-10,13	-15,67	-38,36		0	-12,83	-12,75	-19,5	-6,4	-29,4	-41	12,8
BQG	-4,13	-35,5	-1,8	-16,6	-26,9	-37,2	-36,67		-50	-6,14	-44,5	-10,6	-34	-15,5	-17,6
W1G	34,25	16,67	40,5	17,67	12,43	-15	24,2	19,57		37,17	4,2	21,89	12	-5	-6,13
W2G	16,6	-33,33	-50	-41	-9	-27,25	-14,33	17,09	-36		-3	-17	-14,67	-23,17	-3,2
R1G	37,5	14,08	45	9,63	15,25	-9,5	29,22	14,67	0	-14,5		16,71	-15,44	2,5	-5
R2G	30,6	-33,25	14,17	-20	-20,2	-25	-27,5	15,73	-20,6	-19,33	-30,25		-44	-26,27	-28,6
KAG	50	16,8	24,75	14,2	-14,2	9	0,45	42,78	0,21	36,38	-10	26,6		-15,5	-37
IBG	28,33	10,89	38,67	0,25	18,57	-10,11	15,86	23,6	-0,64	4,88	-8,5	33	5,83		8,4
GHG	7,62	-2,25	13,5	-7,2	-1,6	-6,11	-12	6,2	1,5	-7,29	-18,17	13,88	-13	-7,33	

Defending the average over Ss, we are not interested in interindividual differences but in central tendencies in multidimensional representations based on both preference ratings and perceived similarity which produce and compare, i.e. whether an aggregate solution revealed in the preference data can be reflected in a single configuration retained by aggregating the similarity data. It is also of utmost practical relevance for sound engineering with regard to sensory attributes with reliable scales. To produce a single nonmetric multidimensional solution for a ‘modal person’, the assessments for every pair of noises were averaged over individual data, thus the outcome was a single dissimilarity matrix without diagonals, filled with 210 cells, i.e. 15 % 15 = 225 – 15 (diagonal). The self-dissimilarities lying diagonally were either meaningless or unobserved. Any entry in the matrix had a meaning that one sound sample was preferred over the other. With regard to paired comparisons each subject will ordinarily have clear-cut preferences within each pair of sound samples and that disallows the scaling for any one person. In practice, this problem is to be ignored by pooling the ordinal data averaged over subjects. Such a pooled, averaged dissimilarity data can be useful and

efficient in initial exploration, anyhow at the expense that interindividual differences may cancel out in the averaged group data. Since we had 1200 assessments out of 30 Ss with 40 tasks, the 210 dissimilarities covered six times all possible pairs of 15 sounds, reversals of order included (see Table I-4), i.e. the probability of evaluating all noise pairs at least once is very high. Indeed, every cell in the proximity matrix was assessed by at least five Ss. The nonmetric scaling analysis will break down if there are not enough entries in the dissimilarity matrix to be analyzed. The wavering of a solution is to be released when two thirds of the dissimilarity matrix constitute missing values. Spence (1983) gives evidence for there being a high correlation of 95% between the observed distances and the theoretical ones reconstructed in MDS, even if one-third of the dissimilarities are eliminated, of whatever sort of data (Graef & Spence, 1979).

3.1.2.1 Examination of Symmetry

With regard to the qualitative property of symmetry, Luce et al. (1963, 238) suppose that *"The world is discrete in the sense that there is a finite number of different stimuli, and they are ordered psychologically."*. Symmetry as a requisite for the existence of preference relations appears to be a carry-over from a strictly normative approach to psychological investigation. Symmetry says that there is a single scale, not a collection of unrelated ones.

Table I-5: Symmetry evaluation

p_1Sp_2	p_2Sp_1	*interpretation of symmetry*
+1 (–1)	–1 (+1)	fulfilled
0	0	undetermined
–1 (+1)	–1 (+1)	violated

Of major interest is the case in which the judgment relations are symmetrical, in other words, does the presentation sequence of sounds change merely the sign of assessment? Once the order of presentation does not matter, then one can write $p_1Sp_2 = -p_2Sp_1$. The symmetry evaluation is not metric in nature because from $p_1Sp_2 = 0$ it does not follow $p_1 = p_2$, whereby p_1Sp_2 := symmetry(evaluation(p_1), evaluation(p_2)) holds (see Table I-5). The symmetry evaluation is symmetrical on its own iff $p_1Sp_2 = -p_2Sp_1$, and even linear when the following holds, $(a*p_1)Sp_2 = a(p_1Sp_2)$, and $p_1Sp_2 = p_1S(p_2 + p_3) = p_1Sp_2 + p_1Sp_3$. To achieve a meaningful substantive interpretation of symmetry, above all about its violation, the quantification of symmetry is desirable. To do this,

the proximity data had been converted by $s_{ij} = (p_{ij} \% p_{ji})/(50)^2$ transformations.

Table I-6: Symmetry values obtained from the preference matrix

$s_{ij} = (p_{ij} * p_{ji})/sqr(50)$

$s_i \backslash s_j$	GG	AVG	VMG	ASG	DG	TGG	KG	BQG	W1G	W2G	R1G	R2G	KAG	IBG	GHG
GG		0,62	0,06	0,07	0,44	0,59	0,16	0,00	0,10	0,25	0,70	-0,08	0,93	0,29	0,10
AVG	0,62		0,37	0,02	0,12	0,15	0,24	0,43	0,05	0,29	0,11	-0,09	0,11	0,05	0,00
VMG	0,06	0,37		0,18	0,23	0,31	0,17	-0,02	0,48	-0,51	0,77	0,11	0,41	0,61	0,21
ASG	0,07	0,02	0,18		0,01	0,39	-0,06	0,20	0,26	-0,03	0,04	0,14	0,14	0,00	-0,01
DG	0,44	0,12	0,23	0,01		0,24	-0,12	0,23	0,03	0,07	0,09	0,02	-0,07	0,19	0,00
TGG	0,59	0,15	0,31	0,39	0,24		0,41	0,57	0,08	0,31	0,01	0,28	-0,02	0,03	0,02
KG	0,16	0,24	0,17	–0,06	–0,12	0,41		0,00	0,12	-0,07	0,23	-0,07	0,01	0,26	0,06
BQG	0,00	0,43	–0,02	0,20	0,23	0,57	0,00		0,39	0,04	0,26	0,07	0,58	0,15	0,04
W1G	0,10	0,05	0,48	0,26	0,03	0,08	0,12	0,39		0,54	0,00	0,18	0,00	0,00	0,00
W2G	0,25	0,29	–0,51	–0,03	0,07	0,31	–0,07	0,04	0,54		-0,02	-0,13	0,21	0,05	-0,01
R1G	0,70	0,11	0,77	0,04	0,09	0,01	0,23	0,26	0,00	–0,02		0,20	-0,06	0,01	-0,04
R2G	–0,08	–0,09	0,11	0,14	0,02	0,28	–0,07	0,07	0,18	–0,13	0,20		0,47	0,35	0,16
KAG	0,93	0,11	0,41	0,14	–0,07	–0,02	0,01	0,58	0,00	0,21	–0,06	0,47		0,04	-0,19
IBG	0,29	0,05	0,61	0,00	0,19	0,03	0,26	0,15	0,00	0,05	0,01	0,35	0,04		0,02
GHG	0,10	0,00	0,21	–0,01	0,00	0,02	0,06	0,04	0,00	–0,01	–0,04	0,16	–0,19	0,02	

Thus the value ranges from −1 to +1 (see Table I-6). While checking the violation of symmetry, it follows that symmetry can be retained since the mean of symmetries for all pairs of stimuli revealed +0.16. In total, there were 18 cases concerning the violation of symmetry. Consider an example. The fact that the pair of sounds, (W2G)S(VMG), has the symmetry value of –0.51, can hint at the existence of a position effect.[17] As the average value of symmetry assumes the maintenance of the condition of symmetry, the application of nonmetric scaling analysis to the data set appears justified.

3.1.2.2 Analysis of Dissimilarity Matrix

As Kruskal's analysis considers the monotony condition, i.e. whether the ordering of dissimilarities of the pairs of stimuli is distorted from that of metric distances, the proximities have to be transformed into dissimilarities. Table I-7 is a matrix with entries as metric distances. The nonmetric approach has the mathematical convenience of providing a rational zero, namely, the dissimilarity measure of two identical stimuli. And it is directly associated with the treatment of dissimilarities as metric distances in some latent psychological space. The dissimilarity matrix has been then analyzed with Kruskal's procedure. The calculation of metric distances was done using the Euclidean metric. The distance of a stimulus to itself is null, i.e. the larger the

[17] Stevens named the consistent violation of symmetry as "hysteresis effect" that depends upon factors such as the presentation probability, instructions, payoffs or information feedback (Stevens, 1957).

distance between the two stimuli is the more dissimilar both stimuli would be perceived. By the achievement of an invariant representation of the essential structure, the transformed dissimilarity matrix of the average judgments for 30 Ss on the 210 pairs of sound samples was taken as the input data, yielding quantities that display psychological distances between any two rated samples. Kruskal's analysis was based on the statistical package SYSTAT's MDS module (SYSTAT 9.01, SPSS Inc., 1997), called MDSCAL. MDSCAL contains an adjustable parameter for discontinuing iterations when the value of Stress reaches some predetermined value.

Table I-7: Psychological distances interrelating the 15 sounds used as stimulus-objects in Experiment I

Dissimilarity measure: Euclidean metrics

$d_i \backslash d_j$	GG	AVG	VMG	ASG	DG	TGG	KG	BQG	W1G	W2G	R1G	R2G	KAG	IBG	GHG
GG	0														
AVG	35,87	0													
VMG	22,38	32,38	0												
ASG	28,17	15,89	22,43	0											
DG	30,52	14,58	29,61	15,47	0										
TGG	43,2	16,59	40,98	22,70	19,99	0									
KG	29,37	14,44	27,79	14,88	10,78	23,15	0								
BQG	10,98	31,59	23,18	24,03	26,80	38,28	27,10	0							
W1G	36,81	18,39	33,19	16,75	16,45	14,69	17,82	33,80	0						
W2G	24,99	21,69	29,30	16,77	19,51	28,15	21,35	20,14	26,04	0					
R1G	39,87	15,97	39,44	21,92	14,85	13,82	18,77	35,13	17,41	24,87	0				
R2G	20,74	23,58	17,08	16,35	20,99	30,17	21,24	16,60	24,01	21,29	28,61	0			
KAG	41,14	14,41	39,59	21,20	15,57	12,47	20,01	36,21	16,77	22,69	9,560	28,52	0		
IBG	41,64	18,03	38,71	21,47	16,27	10,71	20,90	37,23	15,04	27,93	15,42	28,49	13,43	0	
GHG	36,08	23,58	34,85	21,49	13,52	23,92	19,67	31,26	20,59	25,21	17,05	26,15	19,46	20,69	0

For analyses of ordinal data the iteration is assessed against the convergence criterion of 0.001 imposed by the MDSCAL. Thus the program would continue to attempt to capture a better solution of stimulus points until either a minimum value of Stress was found, or Stress became 0.001, or the maximal number of iterations was bound to 50. The level of measure was ordinal. The number of points was set at 15 so that there were 105 metric distances. The values of Stress were reported, which were brought about by analyzing in any configuration of points, letting m vary from 1 to 10. Using analytical procedures as described by Kruskal (1964a), it was to be judged whether and how few dimensions were sufficient to represent the psychological space for the preference evaluation.

3.1.2.3 Loss Function Analysis

Stress means the deviation from the perfect correspondence of dissimilarities and corresponding metric distances in the final configuration of stimulus

points. Is Stress zero, then metric distance is a perfect monotonous function of ranked dissimilarity, hence Stress is smaller as monotony increases. The outcome of the analysis of loss function shows the final Stress values relative to different metric coefficients, which is displayed in Table I-8.

Table I-8: Metric coefficient and Stress

Minkowski r-metric	Loss Function
1	0.06064
2	0.03271
3	0.04199
4	0.04588
5	0.04737
6	0.04896
7	0.04998
8	0.05068
9	0.05118
10	0.05153

Stimulus configurations were fitted by scaling with different Minkowski metrics ranging from 1 up to 10. The final configuration was used as the initial configuration that was varied over different metric coefficients. The transition from 2D to 3D-dimensional metric space did not essentially achieve the improvement concerning Stress values. The iteration procedure was stopped after three dimensions to provide a very crude solution. It makes clear that, for dimensionality of three in Minkowski power metric of two, the lowest Stress (.03271) occurs at r = 2, to be interpreted as showing that Stress is acceptably good (see Table I-1), is shown to be fairly well. This means that metric distances are best-fit to the 'true' relations of dissimilarities.

3.1.2.4 Stress Diagram

The goodness-of-fit of the spatial model to the data is a good criterion against which to decide upon the optimal solution of stimulus points. Representing the correspondence of dissimilarities graphically with metric distances in reference to the goodness-of-fit graphically, one gets a sense of the monotonous relation, say Stress diagram, sometimes called Shepard diagram. It provides an overlook of stimulus points scattered around a monotonous representation in which squared metric distances of stimulus points are summed up by a monotonous regression line. It is considered necessary to

capture the monotone mapping of dissimilarities into metric distances in any coordinate space allowing the stimulus points to be fitted in a metric space of minimum dimensionality. The smallest number of dimensions underlying psychological distances in terms of goodness-of-fit appears as the best solution for which one is searching.

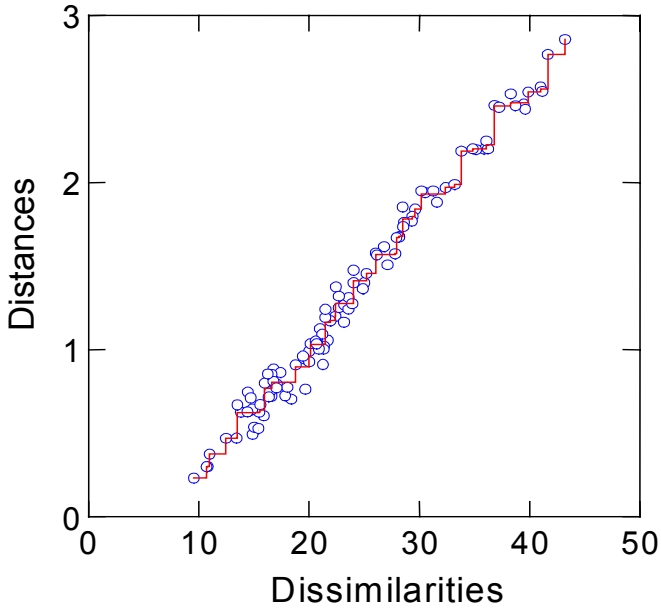

Fig. I-4: Shepard diagram

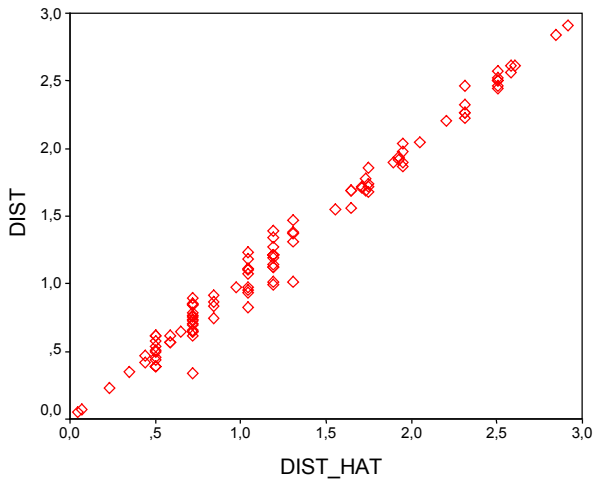

Fig. I-5: Residue diagram

By looking into the Stress diagram, as such is illustrated in Fig. I-4, the scatter of stimulus points does seem to carry a suggestion of an acceptible monotone fit. With Stress being at about 0, all stimulus points should then be on the

bisection line, which means nothing but the correspondence between the geometric and the psychological nearness. A tight coupling appears to exist between the requirements of the formal model imposed by Kruskal's method and the behavior of preference data. This conveys a clue to the fact that the smaller metric distances between two sound samples are, the lower dissimilarities would be. The ordering of dissimilarities in the empirical system is found to agree upon the rank order of metric distances in Euclidean psychological space. For the preference data to which Kruskal's analysis had been applied thus far the monotony condition is shown to be met, hence the suspicion of being degenerate for the final multidimensional configuration can be rejected. On the contrary to the degenerative character of loss function that does cause low Stress, local minimum solutions have an opposite effect. The problem of local minimum to which the non-monotonicity is probably to be attributed might be evoked from the very fact that in scaling experiments, at large, all points of any psychological space could not be realized in the input data. Giving a note that *"The iterative procedures work best when the space is relatively well-filled, but a space with qualitative dimensions cannot be well-filled." (Shepard, 1965, 388)*. Indeed, the upshot can be taken as an indication that local Stress minima proved to be achieved, for the variation of any configurations did not reduce Stress much better. The deviation between the observed distances and the estimated ones is referred to as a residue. Plotting the psychological distances against the estimated ones, it is thus seen that the relationship is relatively linear, as is shown in Fig. I-5. By analyzing residuals based on the dissimilarity values, it is shown that the entire stimulus set could be judged in a fashion consistent with an Euclidean model. Having been the orderings of metric distances tied in with those of dissimilarities, then its disparity is on par with Euclidean distances.

3.1.2.5 Elbow Effect

Empirically deciding on the true metric, Kruskal (1964a) proposes computing different starting configurations for different specifications for fixed pairs of r and m and then selecting the one that leads to the lowest Stress as the true metric. By virtue of the empirical function between Stress and metric coefficients, the optimal spatial model could be determined by putting the Stress values on the ordinate counter metric coefficients laying on the abscissa, letting r and m vary from one through ten in steps of one. If these curves all dip at the same r-value, then it is to be decided that the metric does not depend upon the dimensionality question (e.g. Ahrens, 1974). From Fig. I-6 it can be seen that the lowest Stress occurs at r = 2 over one through five dimensions, demonstrating that subjective distance between stimuli may be

Euclidean.[18] According to Wender (1969), Bortz (1974), Wolfrum (1976), for given 2D configurations, Stress is almost equal for distances with r-exponents of r_1 and $r_2 = r_1/(r_1 - 1)$.

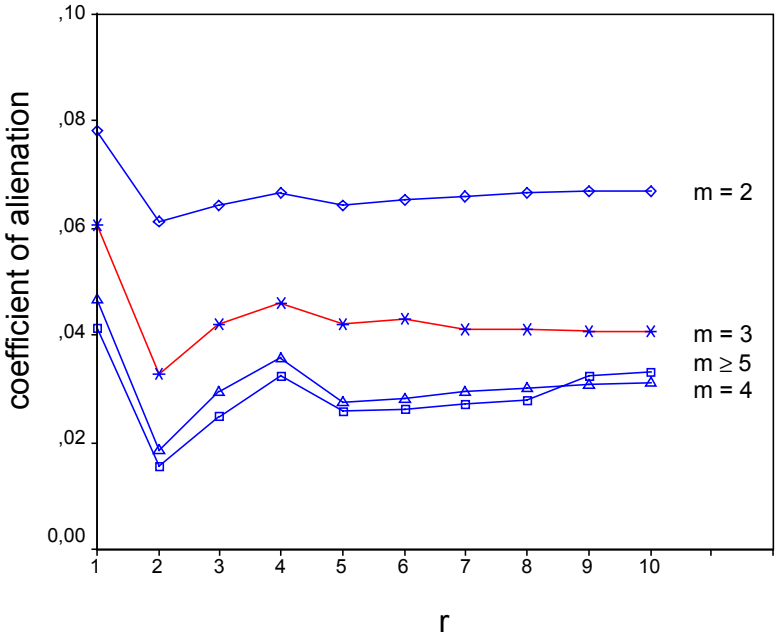

Fig. I-6: s-value relative to r-metrics

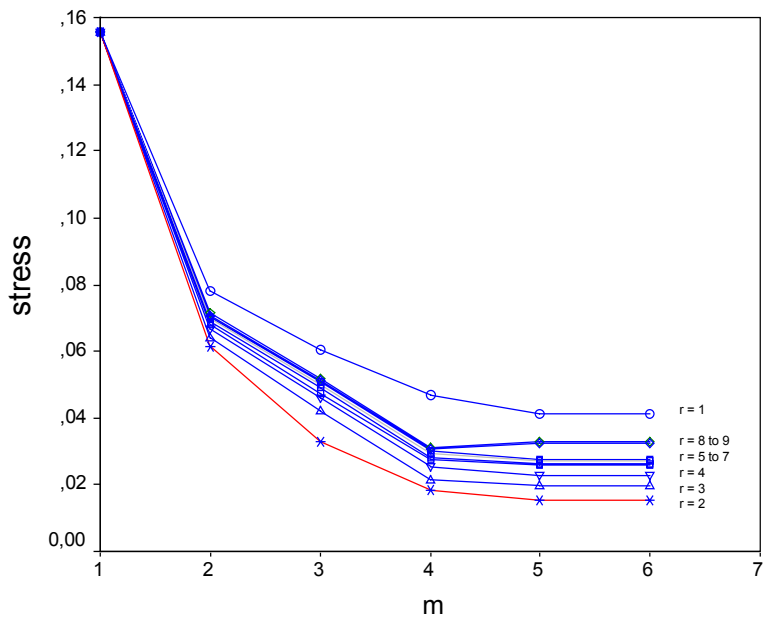

Fig. I-7: s-value vs. dimensionality

[18] To circumvent entrapment in local Stress minima, in general, Kruskal's analysis is initiated in five dimensions (e.g. Wagenaar & Padmos, 1971).

As a general criterion for capturing the appropriate dimensionality of the final configuration of stimulus points, Kruskal (1964a) recommended choosing a value m above which Stress barely changes. To make a decision on the appropriate value of dimensionality for the input data under study, the best solution of stimulus points is constructed in each of several spaces of varying dimensionality. As with the first criterion for selecting the appropriate dimensionality, as was suggested by Kruskal (1964a), the elbow effect, describing a sudden drop of Stress is found to be a diagnostic cirterion for the true dimensionality. The addition of a supplementary dimension improves the Stress marginally. Fig. I-7 shows that the Stress for the input data is a monotonous decreasing function of the number of dimensions. Smooth curves were drawn, through the set of Stress values for each generated dimensionality, hence it was possible to directly read off Stress values for any number of dimensions. The elbow drop on the convex curve track marks the point by which the 'true' dimensionality for the data is deemed to be optimal. A distinctive elbow in the plot of Stress against dimensionality is observable while crossing over from m = 1 to m = 2. Additionally, with reference to the second criterion, as proposed by Kruskal, an (m + 1)-dimensional solution does not contribute to improving the interpretation. Consequently, a 2D solution was accepted as providing the most parsimonious estimates pertaining to the preference structure. The third criterion that considers the statistical variability inherent in the data that was unknown could not be applied in this study.

The problem related to the distribution of Stress has been discussed by Stenson and Knoll (1969), in which the effects of tied ranks were studied. They speculated that the distribution of Stress is quite peaked because the ranges of Stress were 0.02 for m = 1 through 3, 0.01 for m = 4 through 7, and less than 0.01 for m > 7. Accordingly, the preference ratings are decomposable into two dimensions, i.e. the dimensionality chosen for a given configuration of points might be sufficient in accounting for the unique behavior of preference data. The standard deviation of the responses to a particular stimulus is somewhere between 20% and 40% of the mean response value, whilst the errors in physics are usually reduced to less than 1% of the mean. In most social scientific work it is believed that a possible random error of about 25%, i.e. s = 0.25, which real subjects may make in producing the perceptual or cognitive judgments, can be viewed as no exaggeration, especially because elbos effects for Stress values of 0.25 can practically not be observed whatsoever (Discussions on this issue see also Wagenaar & Padmos, 1971).

3.1.2.6 Spatial Representation

The next step was to derive a multidimensional set of stimulus coordinates from the data without reference to the component dimensions. Since the third dimension accounts for so little, we may simply ignore it. MDS geometry serves as a model of psychological space. By a closer look at the image of some underlying composition rule for the basic dimensions of the sound samples, which is displayed in Fig. I-8, we find that the preference judgments are substantially oriented in two dimensions along which all sound sources relative to each other could be ordered. The orthogonality of principal axes fixed in a scaling space means that perceptual dimensions are empirically subjected to be independent. At first glance we see a simplex organization pattern indicating that sound samples are almost scattered along Dimension I. A viable alternative to attain the substantive interpretation of the solution with regard to the factors explaining the ordering of stimulus points along the perceptual dimensions and to generalize the irregular simplex organization pattern involves a multivariate technique, namely the clustering or tree procedures that have a nondimensional character (e.g. Grey, 1977).

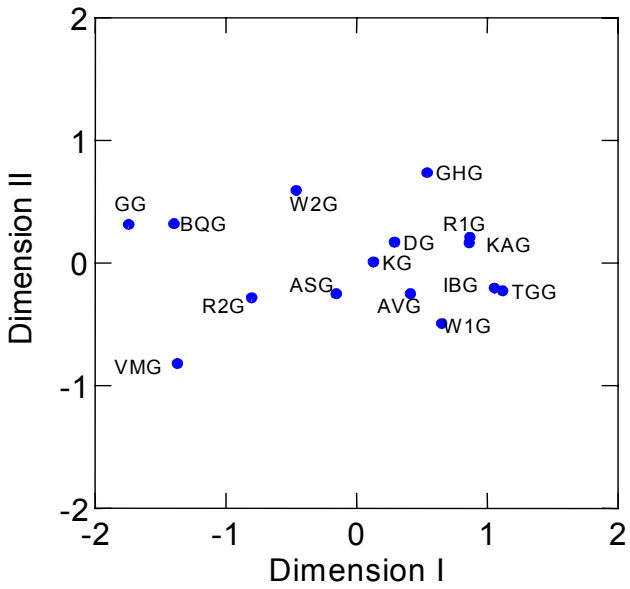

Fig. I-8: NMDS-configuration

Indeed, Shepard (1972, 1974) proposed that MDS combined with the clustering procedures may lead to the substantive interpretation of sensory attributes pertaining to the perceptual data. The clustering procedures such as hierarchical linkage algorithm (McQuitty, 1960) or hierarchical cluster algorithm ("HICLUS", Johnson, 1967), invariant for monotone mapping of dissimilarities into metric distances in a coordinate space, provide additional information about what becomes not visible from merely spatial representation. The hierarchical cluster procedure implemented to ferret out

homogeneous groups in the preference data revealed a hiearchical system of clustering representations in which all stimuli were grouped together as a single cluster. Hence, to derive a more reliable number of clusterings and to model dissimilarities, the additive tree procedure advanced by Sattah and Tversky (1977) was performed. The stimuli belonging to a cluster can be compared by their node lengths.

Table I-9: Calculation of node lengths between the 15 sounds based on the additive trees analysis in Experiment I

Node	Length	Node	Length	Cluster
GG	5.370	BQG	1.567	5.814
TGG	3.830	IBG	2.839	2.420
R1G	2.824	KAG	2.696	1.232
VMG	9.455		5.814	2.093
	2.420		1.232	2.326
	2.093	R2G	2.666	4.885
DG	2.099	GHG	7.377	0.395
W1G	4.740		2.326	1.458
	4.885	W2G	5.769	2.286
AVG	4.471		1.458	1.861
ASG	2.768		2.286	0.000
	1.861		0.395	1.796
KG	4.016		1.796	2.820
	0.000		2.820	0.000

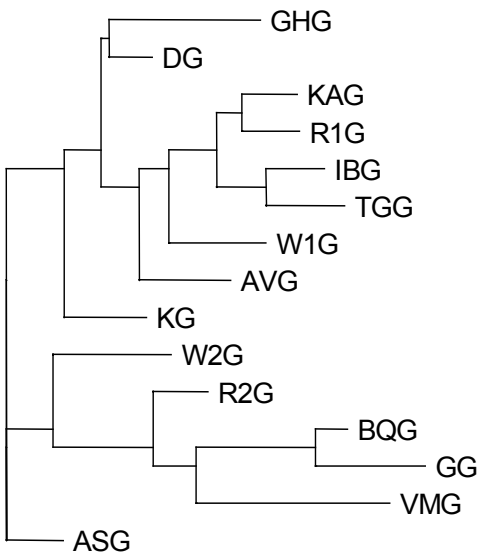

Fig. I-9: Additive trees

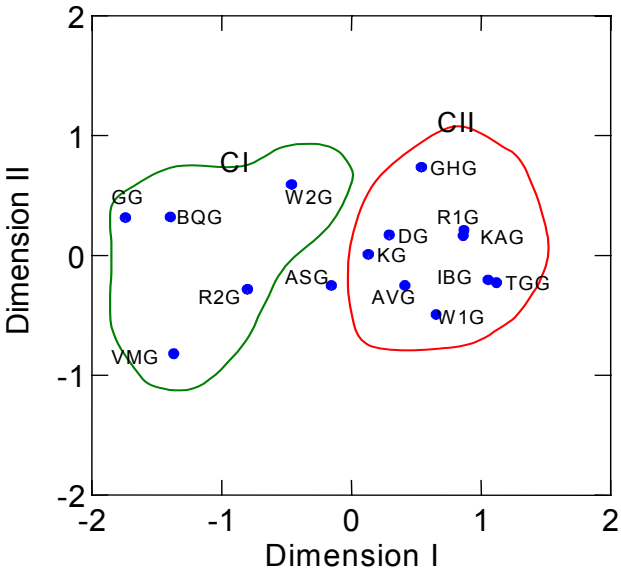

Fig. I-10: Stimulus groupings

The additive tree results confirmed the discontinuous categorical structure of stimuli. The representation of sound similarity as an additive feature tree, which is displayed in Fig. I-9 and listed in Table I-9, is indicated by the contours of grouping sounds that belong to the same class. There were two dominant clusters CI, CII with the exception of ASG that fit the pattern of the scaling space but do not belong to any established cluster. This can be easily seen in Fig. I-10. Hiearchically sorting by nodal pairs, with cross-linkage permitted, and sorting by ranking all pairs, the algorithm of the additive tree has ended with two groupings of stimuli. Additionally, this additive tree checks on the condition of ultrametric inequality, in terms of whether the distance between two stimuli is not larger than its distance relative to a third stimulus at all. The axiom of ultrametric inequality proved to be fullfilled. The distance between two nodes is the sum of horizontal lengths of both. Consider an example. The distance between the nodes VMG and GG is 9.455 + 5.814 + 5.370 = 20.639, the distance between the nodes VMG and BQG is 9.455 + 5.814 + 1.567 = 16.836. From this one can draw a conclusion that VMG is more similar to BQG than to GG. On the other hand, it was observed that the groupings retained from this multivariate analysis do not correspond to the families of sound sources classified at the outset. Two classes of sounds are distinguishing: Inspection of Fig. I-10 suggests that Dimension I contrasts the stimuli aligned with W1G, TGG, IBG, KAG, KG, DG, R1G, GHG, and AVG, all of which are relatively low and soft in character, with those aligned with VMG, GG, R2G, W2G, and BQG, which are heard as accelerating and increasingly loud. On the other hand, Dimension II contrasts the stimuli aligned with GG, BQG, W2G, and GHG, all of which have more or less irregular or disrhythmic characteristics, with VMG and R2G, both of which are distinguished as relatively sharp and high. Those sounds like VMG, GG, BQG, R2G, and W2G belonging to CI are characterizable as loud and rhythmic compared to those aligned with W1G, TGG, IBG, KAG, KG, DG, R1G, GHG, and AVG in CII, all of which are heard as relatively soft and low. Altogether three dimensions accounted for 82.3% of the variance in the data. The first dimension explained 65.7%, the second dimension 9.6%, and the third dimension accounted for 7% of the variance in the data. The best contribution to predictions is made by the first dimension. The information gain brought about by going from the first to the second or to the third dimension seems to be negligible since the explained variance is very low.

3.1.3 Discussion

Addition of the third dimension did not increase the interpretability so much as it being worth choosing a three-dimensional solution for the sake of representing the psychological space for preference judgments. Generally speaking, increasing preference appears to run according to the absolute value of the stimulus projection on the first dimension. One of the relations which could not be established by the analysis was the extent to which generally conceived preferences influence the personal preference evaluation. The preference relations among the positive stimuli on the first dimension and the negative ones on this dimension are difficult to assess in isolation from the utilization of the space by the individual. Approaching the positive side on the dimension should not be considered an indicator of liking. It is felt that when judging the relative degree of preference the Ss tended to discriminate the stimuli essentially with respect to a specific sensory attribute. Ss reported from questionnaires that they rated all stimulus pairs in terms of such descriptive adjectives like quiet, smooth, windy, regular, dull etc. for preferred sounds, those like loud, accelerating, high, shrill etc. for disfavored ones. Semantic descriptors rather describe reactions to the stimuli than the nature of stimuli. Dimension I, in terms of attributes which might be termed strength is deemed to be a substantial property for subjective preference of the sounds compared to an additional Dimension II that has to do with semantic descriptors like speedy, rhythmic, accelerating, etc. To summarize the discussion up to this point it may be said that the Ss seem to attend mainly to an apparently conspicuous attribute of sounds which is reflected in the multidimensional representation of stimulus points. Preference scaling generated a metric space dominated by Dimension I, while Dimension II plays a much less crucial role.

Along with Zajonc (1980, 159), it should be noted that the analysis of preferences in the sense of “hot cognitions” coined by Abelson (1963) is not simply an analysis of a spatial representation in the sense of “cold cognitions” that have become hot. Nonmetric scaling has a possibility not only to investigate an explicit hypothesis but also the dimensionality and the metric by means of data. It is very likely that a 2D Euclidean model concerning the conditions of monotony and economy is to be selected as an optimal rule of cognitive combination for sound preferences. As Euclidean transformations would not affect the metric distances, dissimilarities consistent with metric distances in an Euclidean space can be represented and suitably rotated or stretched until the principal axes have a plausible meaning in terms of stimulus properties. However, a method of nonmetric scaling which assumes that a spatial model is an appropriate representation for the data is not legitimate at all because multidimensional scaling could lead to plausible but spurious stimulus dimensions for complex stimuli (e.g. Cliff, 1969). As is commented earlier by Beals and Krantz (1967), there is no cogent argument to treat the preference ratings in terms of a geometric representation like Minkowski spaces. Despite these concerns the Euclidean model has served us

well not as a description of preference relations but rather as a baseline against which to compare patterns of deviations. As far as one does not find himself in the position to a priori justify explicit hypotheses concerning configuration, metric, and dimensionality of data, solutions of principal axes and its interpretation should be seen merely as a hypothesis. On the grounds of results of the loss function analysis, the retained perceptual dimensions could serve for further developing the hypothesis.
By using reproductions of paintings, Berlyne and Ogilvie (1974) reported that similarity judgments can be represented in a three-dimensional space, and that the judgments of preference and similarity are subjected to the same three perceptual dimensions, labelled Hedonic Tone, Arousal, and Uncertainty, respectively. One further point concerns the stability of the derived structure, i.e. whether it remains robust even when the Ss are required to directly assess similarity of sounds. By performing the nonmetric analysis using a separate proximity matrix for both preference ratings and perceptual similarities, it would be interesting to compare the coordinates of stimulus points in a space of prespecified dimensionality. In doing so, it will be shown which perceptual dimensions will have differential saliences or weights, i.e. specific cognitive bias in different judgment types. Discussions related to the issue of interindividual differences in multidimensional scaling can be found in the papers by McGee (1968), by Horan (1969), and by Carroll and Chang (1970). Viewing the multidimensional unfolding model (e.g. Bennett & Hays, 1960; Carroll & Chang, 1970), the subjects agree on the structure of the similarity space and only disagree on their preferences among the pairs of stimuli. Positive correlations between unidimensional preference scales and a multidimensional similarity space have been reported by Cliff and Young (1968), by Isaac (1970), and by Steinheiser (1970). Therefore, we turn now to a second experiment in order to see whether the subjective attributes revealed in pairwise preference ratings will be reflected in those pertaining to the similarity evaluation.

3.2 Experiment II : Similarity Scaling [19]

In this section of a series of experiments we are interested specifically in the possible compatibilities of the results comparable to those obtained from the preference study. In order to get an idea of the contribution of the perceptual dimensions revealed in the preference study to the perceived similarity, a listening experiment was conducted with 23 subjects. The stimuli were the same as those used in the preference study. The technique used to elicit pairwise similarity judgments was effectively equivalent to that applied in the

[19] This similarity study is presented as part of technical report at the Institute for Research into Man-Environment-Relations, Oldenburg, submitted in November 2000 (Choe, 2000b).

foregoing experiment. A single dissimilarity matrix without diagonal was computed. 210 entries in the matrix were analyzed using Kruskal's procedure, producing a solution configuration. We started the hypothesis assuming that the perceptual dimensions pertaining to the preference evaluation of vehicle noises could be reflected by underlying the similarity evaluation on the same set of sound sources. It will be shown whether the perceptual criteria by which pairwise preference ratings were made would be reflected in the solution retained in the similarity structure. Pairwise similarity judgments provide a valuable tool in the study of perception and cognition. Similarity experiments are generally employed when the stimuli have no known ordering other than a subjective one.
As has been discussed by McAdams et al. (1995), multidimensional scaling of similarity judgments has the advantage in its relative ease of assignment and no need of a priori assumptions regarding the number of cognitive factors underlying the geometric representation. One month later [20] the 23 individuals recruited from Ss who took part in Experiment I also served as Ss in the similarity study. The group had seven Ss missing, 12 females and 11 males, aged 28 years on average. Ss were instructed to assess the degree of overall similarity toward sound sources on a bipolar rating scale with interval character. By instructing subjects to rate the degree of overall similarity for each pair of sound sources, all Ss adequately completed all tasks and were included in the analysis.

3.2.1 Method

A program module implemented on the NOVAL interface was adequately configurated for the purpose of assessing pairwise similarity judgments for pairs of sound samples (see Fig. II-1). Except as otherwise specified, Experiment II was conducted using the pattern of Experiment I. A set of sound samples was representative of vehicle noises we are normally exposed to. 40 paired comparisons of 15 vehicle interior noises were presented, identical to those in the foregoing experiment. The Ss listened to the 40 possible pairs via Sennheiser headphone HD 540 reference II in a hemi-anechoic chamber. One alteration was made in a 100 step rating scale. Ss made the similarity judgment by moving a track bar on the program interface that represented a continuous bipolar rating scale from unequivocal dissimilar (coded –50) to unequivocal similar (coded +50), marked with verbal labels at the endpoints and no labels in between. When the perceived value between two sound samples being compared increases and is positive, then it can be stated that the first sound is felt to be similar to the second sound. The same

[20] Having performed pairwise evaluations of both preferences and similarities subsequently, Cooper (1973) found that, when preferences collected first, similarities were strongly biased. In his study, however, a minimum of one hour separated consecutive sessions.

holds for the negative case. Once a first coarse adjustment is achieved, the S can proceed more subtlety and with precision. Prior to conducting the main session, 3 practice trials were run in order to familiarize the Ss with the task. Each sound lasted for 15 seconds. As with the main trials, each S was given the same order of presentation of the paired samples as had been established in the foregoing experiment. Each sound pair could be played back on request. Once the judgment was entered, the next trial was automatically presented. Each session lasted about 45 min.

Fig. II-1: Program interface for similarity evaluation

Block II: Instruction for the similarity-difference judgments in Experiment II

Thank you for your readiness to take part in my experiment. The experiment will take 45 minutes. In the following test you will hear two sounds from cars. Your task is to compare the two heard sounds. Please proceed as follows:
Click on the left button for the first sound with the left button of the mouse, and listen to it. After having heard the first sound, click on the right button for the second sound and listen again. You listen to every

> sound until it ends. Notice that you don't start the second sound before the first sound has ended. It should be noted that certain sounds in combination with other different sounds could occur. Every sound lasts 15 seconds. Each time a pair of sounds is displayed and heard, you are to decide how much alike or different they seem to you, and rate them accordingly on the rating scale provided. To judge sounds a rating scale is provided, which has a horizontal track bar. The rating scale at hand has 2 labels at the ends:
>
> ***- I think the both sounds to be unmistakenly dissimilar.***
> ***- I think the both sounds to be unmistakenly similar.***
>
> You can move the track bar on the whole scale with the mouse. To move the track bar, push it by draging the left button of the mouse, and set it to your desired position on the scale, or for adjustment you can set down the track bar somewhere between the extremes while pressing on the range of scale. At the beginning of each session the track bar is automatically set in the middle. When you are not able to decide after having heard a pair of sounds, you can repeat each sound once again. Having made a decision, press the button "confirm". After that, you obtain a message "Your decision was confirmed!". Please press the buttton "OK", then the next pair of sounds is ready. The track bar is back to the middle of the scale. Just make your ratings on the basis of their overall similarity, rather than on any other particular characteristic. The first two ratings you make will be just for practice. This practice series will familarize you with the sounds you are to rate. If having difficulties dealing with the task, or if lacking claritiy about the course of experiment, you can either retrieve the instruction by pressing button "instruction", or you may ask the experimentator at any time. Shortly after having solved the last pair of sounds, you will obtain a message of the end of the experiment. Then you may return to the experimentator.

It was stressed that the task merely relies on overall impressions that immediatly spring to subjects' mind. After completing the listening task, a questionnaire followed. Ss were required to name descriptive adjectives in terms of individual's judgment criteria that served the similarity assessments. The instruction button is located on the upper side of the program interface, and instructions contain a few minor revisions (see Block II).

3.2.2 Treatment of Data and Results

Identical analyses applied to the preference case were performed on the similarity data. To begin with, a matrix of proportions of similarities was computed on the whole set of listeners. Omitting the practice series from consideration, a total 920 similarities were obtained. Aggregation over Ss for all similarities produced 210 mean similarity values. An entry in the

proximity matrix thus means that two sound samples being compared were similar or different to each other. Interstimulus distances were calculated using nonmetric scaling analysis. The 210 dissimilarities covered five times all possible pairs of 15 sounds, reversals of order included (see Table II-1).

Table II-1: Scale values for paired sounds based on ratings on similarity-difference continuum

$s_i \backslash s_j$	GG	AVG	VMG	ASG	DG	TGG	KG	BQG	W1G	W2G	R1G	R2G	KAG	IBG	GHG
GG		-48,50	44,67	-2,33	-3,00	7,50	-19,40	-44,40	-50,00	-22,80	-30,50	-31,50	-25,67	-8,00	-42,00
AVG	-39,50		-31,40	28,67	-17,00	24,71	23,00	-50,00	27,67	-4,00	0,33	-45,50	11,00	1,55	-36,33
VMG	29,71	-41,00		3,50	6,00	-4,40	-15,67	-26,80	-34,00	-6,00	-44,17	-39,75	-41,00	-25,75	-48,67
ASG	5,83	9,60	24,50		6,33	5,38	7,50	-40,40	13,00	9,88	28,00	-36,75	-7,75	-23,50	-24,80
DG	26,00	-39,40	7,00	-9,33		-27,67	-19,67	-43,29	-20,71	-20,25	0	-33,57	-18,43	13,20	-33,20
TGG	-27,00	25,50	-8,00	13,67	-3,00		16,00	-45,00	14,83	-13,20	24,25	-40,67	41,00	-4,71	-30,20
KG	-4,50	5,33	-3,33	28,33	5,00	7,00		-50,00	-9,50	9,50	-5,67	-41,50	-2,00	-25,00	-38,00
BQG	-44,00	-44,75	-45,33	-37,50	-35,44	-45,00	-47,20		-16,50	-43,20	-50,00	-43,20	-45,00	-45,20	-13,00
W1G	-41,67	12,20	-41,50	0,75	-30,67	5,00	-0,82	-39,00		-2,20	15,00	-33,86	33,00	-13,00	-7,80
W2G	-17,00	8,50	0	15,00	6,33	2,25	-34,75	-39,88	-2,50		18,50	-30,20	12,20	-5,20	-32,60
R1G	-29,80	6,91	-48,50	-1,25	-50,00	32,33	-0,71	-38,50	0	6,50		-35,50	37,25	-1,25	8,67
R2G	-47,50	-46,00	-47,00	-44,00	0	0	-23,50	-23,89	-25,50	-42,50	-50,00		-46,50	-26,13	-38,80
KAG	-29,50	23,86	-20,00	9,40	-38,25	18,67	18,75	-45,29	26,83	11,00	39,33	-34,75		0,00	-1,50
IBG	-6,00	-5,38	-26,71	-2,25	3,33	8,00	-10,00	-46,25	-22,64	-29,33	13,50	-44,22	-8,67		-50,00
GHG	-46,08	-47,50	-48,25	-46,33	-44,50	-30,43	-44,50	-2,20	-37,00	-49,50	-28,20	-30,00	-41,50	-31,00	

Indeed, every sound pair was rated by at least four Ss. A single set of dissimilarities thus produced was analyzed with Kruskal's procedure. To advance the Minkowski power metrics as a strong model of perceptual judgments is to assume that the metric conditions like identity, symmetry, and ultrametric inequality, be met.

Table II-2: Symmetry values obtained from similarity matrix

$s_{ij} = (a_{ij} * a_{ji})/sqr(50)$

$s_i \backslash s_j$	GG	AVG	VMG	ASG	DG	TGG	KG	BQG	W1G	W2G	R1G	R2G	KAG	IBG	GHG
GG		-0,77	-0,53	0,01	0,03	0,08	-0,03	-0,78	-0,83	-0,16	-0,36	-0,60	-0,30	-0,02	-0,77
AVG	-0,77		-0,51	-0,11	-0,27	-0,25	-0,05	-0,90	-0,14	0,01	0,00	-0,84	-0,10	0,00	-0,69
VMG	-0,53	-0,51		-0,03	-0,02	-0,01	-0,02	-0,49	-0,56	0,00	-0,86	-0,75	-0,33	-0,28	-0,94
ASG	0,01	-0,11	-0,03		0,02	-0,03	-0,09	-0,61	0,00	-0,06	0,01	-0,65	0,03	-0,02	-0,46
DG	0,03	-0,27	-0,02	0,02		-0,03	0,04	-0,61	-0,25	0,05	0,00	0,00	-0,28	-0,02	-0,59
TGG	0,08	-0,25	-0,01	-0,03	-0,03		-0,04	-0,81	-0,03	0,01	-0,31	0,00	-0,31	0,02	-0,37
KG	-0,03	-0,05	-0,02	-0,09	0,04	-0,04		-0,94	0,00	0,13	0,00	-0,39	0,02	-0,10	-0,68
BQG	-0,78	-0,90	-0,49	-0,61	-0,61	-0,81	-0,94		-0,26	-0,69	-0,77	-0,41	-0,82	-0,84	-0,01
W1G	-0,83	-0,14	-0,56	0,00	-0,25	-0,03	0,00	-0,26		0,00	0,00	-0,35	-0,35	-0,12	-0,12
W2G	-0,16	0,01	0,00	-0,06	0,05	0,01	0,13	-0,69	0,00		-0,05	-0,51	-0,05	-0,06	-0,65
R1G	-0,36	0,00	-0,86	0,01	0,00	-0,31	0,00	-0,77	0,00	-0,05		-0,71	-0,59	0,01	0,10
R2G	-0,60	-0,84	-0,75	-0,65	0,00	0,00	-0,39	-0,41	-0,35	-0,51	-0,71		-0,65	-0,46	-0,47
KAG	-0,30	-0,10	-0,33	0,03	-0,28	-0,31	0,02	-0,82	-0,35	-0,05	-0,59	-0,65		0,00	-0,02
IBG	-0,02	0,00	-0,28	-0,02	-0,02	0,02	-0,10	-0,84	-0,12	-0,06	0,01	-0,46	0,00		-0,62
GHG	-0,77	-0,69	-0,94	-0,46	-0,59	-0,37	-0,68	-0,01	-0,12	-0,65	0,10	-0,47	-0,02	-0,62	

Consider an example. The pair of sounds, (W2G)S(KG), which has a perceived value –34.75 on its own, is in disagreement with its counterpart, (KG)S(W2G), yielding a perceived value of 9.50. Likewise the pair, (TGG)S(GG), which has a perceived value of –27, seemingly is not symmetrical to its counterpart, (GG)S(TGG), having a perceived value of 7.5. The value of –50 (or +50) implies that both stimuli were judged as being different (or similar), and the value of 0 implies that both stimuli were not comparable. In total, there are 15 cases in which the symmetry condition is proved violated (see Table II-2). Of course, it should be noted that the quantification of symmetry, as was done in Experiment I, has a different meaning when applied to the present data. In view of the fact that a symmetry quantity does not enable us to make decisions about which of two stimuli is laid down as a leverage that is qualified as being more different or similar to the compared one, hence the average value of symmetries as –0.27 suggests merely that Ss had a tendency to perform tasks on the basis of differences rather than on the basis of similarities.[21]

3.2.2.1 Analysis of Similarity Response

The nonmetric multidimensional scaling technique is inherently metric and gives a multidimensional configuration of stimuli which is based on the data from Ss in the sample. The similarity judgments made by Ss were converted to dissimilarities. Table II-3 is a matrix whose entries are metric distances, which were analyzed with Kruskal's procedure. The calculation of metric distances was done using the Euclidean metric.

Table II-3: Psychological distances interrelating the 15 sounds used as stimulus-objects in Experiment II

Dissimilarity measure: Euclidean metrics

$d_i \backslash d_j$	GG	AVG	VMG	ASG	DG	TGG	KG	BQG	W1G	W2G	R1G	R2G	KAG	IBG	GHG
GG	0														
AVG	39,12	0													
VMG	13,31	39,87	0												
ASG	33,44	21,84	32,03	0											
DG	20,19	35,44	22,57	28,55	0										
TGG	38,95	24,78	39,35	19,04	32,32	0									
KG	34,71	18,89	35,23	17,42	30,16	18,59	0								
BQG	36,79	45,39	42,91	50,28	37,27	49,22	41,85	0							
W1G	37,67	14,01	39,12	23,44	33,89	25,04	16,63	43,46	0						
W2G	27,98	19,08	30,40	16,69	28,05	21,79	14,09	39,27	19,96	0					
R1G	39,99	16,33	36,02	23,38	35,34	25,43	23,79	51,71	20,92	23,85	0				
R2G	32,57	38,94	34,99	43,91	31,93	43,51	36,55	10,99	38,75	32,94	45,21	0			
KAG	44,11	15,12	43,19	23,61	39,02	22,25	23,23	50,19	19,94	24,08	15,51	44,28	0		
IBG	27,78	27,31	28,94	22,55	25,44	23,69	19,07	34,93	23,55	19,70	25,39	28,19	26,21	0	
GHG	37,58	30,49	39,94	40,08	39,35	38,57	30,43	22,85	29,77	25,86	38,95	20,05	35,86	25,06	0

[21] Following Torgerson (1965), an adequate and meaningful spatial representation can be found even though the empirical data fails to satisfy certain necessary conditions like symmetry.

A psychological space with two perceptual dimensions was to be captured. The adopted iteration procedure imposed by Kruskal's procedure was applied to the similarity data as usual, having the following properties: the level of measure was ordinal; the iteration was stopped when Stress converged to a value of 0.001; the maximal number of iterations was set to 50. The appropriate spatial model taken causes the lowest residual sum of squares. The final Stress values related to different metric coefficients are shown in Table II-4, fitted by choosing different Minkowski r-metrics from 1 up to 10.

Table II-4: Metric coefficient and Stress

Minkowski r-metric	Loss Function
1	0.07595
2	0.02868
3	0.02915
4	0.03328
5	0.03641
6	0.03847
7	0.04043
8	0.04204
9	0.04334
10	0.04441

The iteration procedure was stopped after three dimensions to provide a very crude solution. The Stress computed as 0.02868 for 3D in a metric parameter of two was decided upon since the transition from 2D to 3D-dimensional metric space does not improve Stress substantially. When the formal requirements imposed by nonmetric scaling that the orderings of dissimilarities should be linearly related to those of metric distances in any dimensionally organized space would not be violated, one can in fact use similarity to capture the underlying structure of the stimuli under study. The nonmetric scaling algorithm reports a goodness-of-fit statistic between the subjectively judged distances of the stimuli and their reconstructed metric distances. This can be seen in Fig. II-2, thereby representing no distortion between the ordering of dissimilarities and that of metric distances, yielding an acceptible monotone fit. The stepwise or L-formed Stress function might well be hinting at the existence of a degenerate solution for the final configuration. The problem of degeneracy, sometimes called nonuniqueness, is referred to as an undesirable loss or an extraneous imposition of structural information (Shepard, 1974). In general, a degenerate or non-unique solution

occurs as a result that the number of stimuli is shown to be smaller than that of dimensions (Shepard, 1966). It should be noted, as already pointed out by Gregson (1975), that it is possible, by making a wrong assumption about the metric properties of dissimilarities on a set of stimuli of unknown composition, to produce a multidimensional configuration with too few dimensions correctly identified and some additional pseudo-dimensions found.

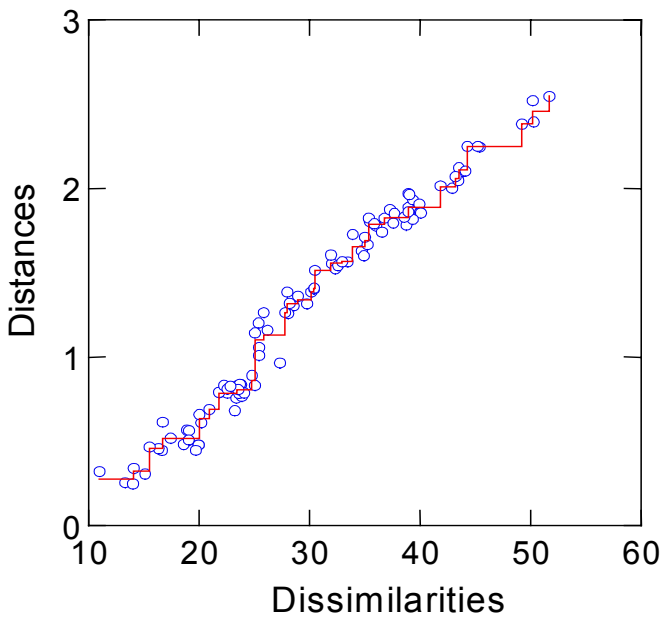

Fig. II-2: Shepard diagram

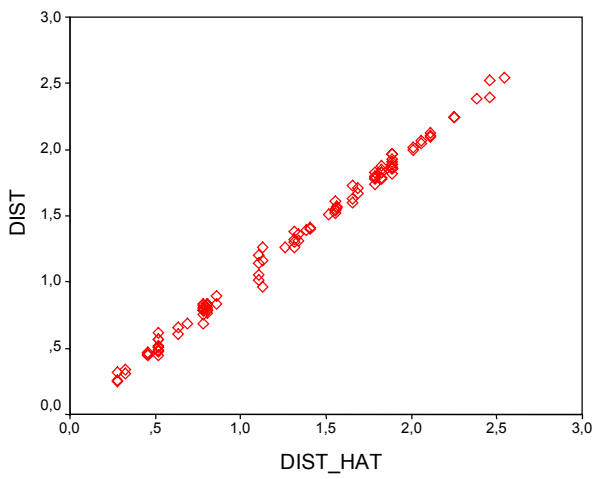

Fig. II-3: Residue diagram

On the other hand, the minimal dimensionality criterion imposed by the nonmetric scaling algorithm can by itself conflict with the requirements of a

psychological model in a particular context (e.g. sound evaluation in field can ask for more dimensions than does in lab.), or how the investigator thinks of how the psychological scaling ought to look if it is meaningful. Suppose, if enough stimulus points and not too many dimensions, nonmetric scaling does work, i.e. any effects of monotonous bias in metric distances result in an increase in dimensionality. During the process of minimalization Stress can be diminished even though dissimilarities and metric distances do not stand in systematic relationship with each other. If this is actually the case, we infer that the loss function might degenerate, albeit Stress being low, the monotony condition is violated, i.e. an ordering of dissimilarities is not compatible to that of unequivocally metric distances. This suspicion runs counter to the finding that the monotone fit increases, i.e. the monotony condition proves to be fullfilled. The failure to obtain a perfect correspondence between the data and the representation seems to be attributable to a large number of ties in the data. Where tied ranks happen to occur, the goodness-of-fit measure can be lowered. As a matter of fact, Stenson and Knoll (1969) reported that 10% of ties can reduce Stress by a relatively small amount of less than 0.02 when m > 3. Their input data were 190 two-digit random numbers, which were analyzed in a given dimensionality, letting m vary from 1 through 10 in steps of 1. In the present sutdy, we found 10 ties total in the ordering of 210 metric distances for 15 sound samples, which amounts to the effect of 4.76% of tied ranks. Hence, it is to be concluded that the occurence of ties did not significantly affect Stress. Other possible reasons have already been discussed in section 3.1.2. Having undertaken an analysis of residuals on the judgments of 23 Ss to 15 stimuli, as is shown in Fig. II-3, we can end up with an Euclidean model that can fit the data very well.

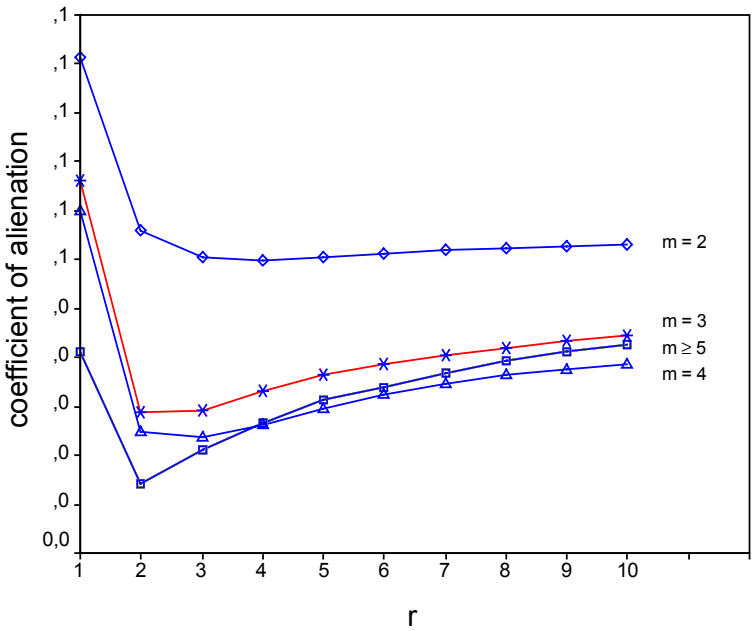

Fig. II-4: s-value relative to r-metrics

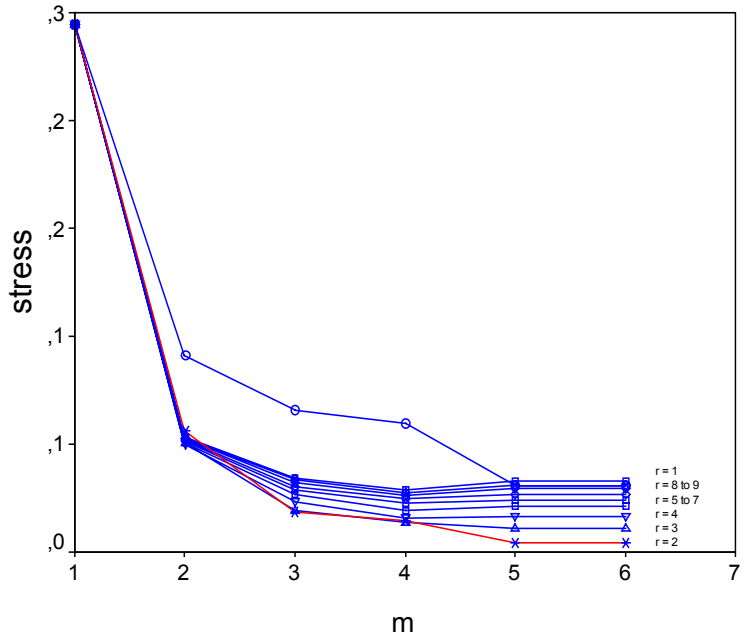

Fig. II-5: s-value vs. dimensionality

Out of Fig. II-4, the finding that the lowest Stress occurs at r = 2 for three or four dimensions, at r = 3 for two or over five dimensions, may be evidence that the subjects used an r parameter close to 2 (e.g. Arabie, 1991). With more than three dimensions, the Stress could hardly be improved. We then go a step further in order to capture the underlying dimensionality pertaining to the similarity judgments. Also shown are the average values of Stress as a function of the number of dimensions using the Euclidean metric. As is displayed in Fig. II-5, there is a quite distinctive elbow in the plot of Stress against dimensionality, i.e., a major decrease in the marginal improvement effected by an additional dimension used essentially only to scale the noise in the data (e.g. Borg & Groenen, 1997, 38). The dimensionality of two is thus to be determined in that Stress is acceptably small, just for reasons that an (m + 1)-dimensional solution adds no further information for interpreting MDS space ("second criterion" according to Kruskal, 1964a). As a consequence, 2D solution was accepted as providing the optimal estimates retained in the similarity structure, although marginally open to the possibility of an argument in favour of three.

3.2.2.2 Dimensional vs. Nondimensional Interpretation

The next step was to derive a multidimensional set of stimulus coordinates from the distance data without reference to the component dimensions. Even

when an Euclidean spatial model turns out to be an appropriate representation for the psychological structure of a given set of stimuli, the true importance of perceptual dimensions pertaining to the similarity judgments has to be investigated. To capture the sensory attributes of similarity judgments for the panel of listeners, we now turn to a plot of the scaling space for the similarity data. As was the case in the preference data, the three-dimensional solution in the similarity data was not directly interpretable. By means of Stress values, there is a very clear 2D solution for the scaling space for similarities, which accounts for 71.3% of the variance of the values. Because the addition of a third dimension did not improve the interpretability, we concentrate on the plane spanned by the first two dimensions, which is presented in Fig. II-6.

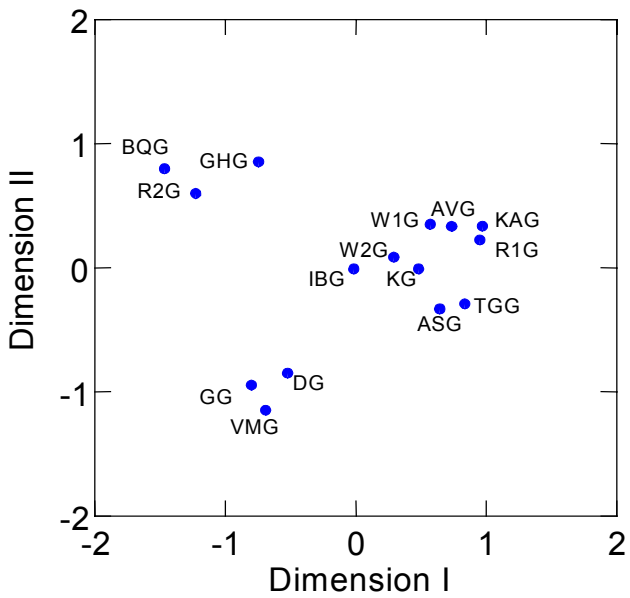

Fig. II-6: NMDS-configuration

We see an irregular circumplex pattern, hence a further multivariate analysis seems to be demanded. As usual, we undertook a additive tree procedure to classify the stimuli into optimally homogeneous groups hidden in the given array of numerical data. This is shown in Fig.II-7 and listed in Table II-5. The axiom of ultrametric inequality proved to be met. From the representation of sound similarity as an additive feature tree, three distinct clusters were found. This is illustrated in Fig. II-8. Dimension I contrasts the cluster CIII that contains sounds aligned with ASG, TGG, KG, IBG, W2G, W1G, R1G, KAG, and AVG, all of which are relatively low and soft in character, with the clusters CI, CII including those aligned with BQG, R2G, GHG, GG, VMG, and DG, which are heard as accelerating and increasingly loud. CI and CII are apparently separated from one another by adding the second dimension. Dimension II contrasts CI involving BQG, R2G, and GHG, characteristics of

which are considered to be more or less discontinuous or disrhythmic in nature, with CII that has those sounds like GG, VMG, and DG, all of which are heard relatively as sharp and high.

Table II-5: Calculation of node lengths between the 15 sounds based on the additive trees analysis in Experiment II

Node	Length	Node	Length	Cluster
GG	6.127	VMG	7.180	5.688
	8.352	R2G	2.639	8.377
R1G	7.572	KAG	7.941	1.681
	5.688	DG	9.042	5.101
AVG	6.286		1.681	2.158
	8.377	GHG	7.574	7.490
	5.101		2.913	1.431
	2.158	W1G	7.744	2.652
ASG	8.506	TGG	10.29	1.706
	4.603	IBG	6.960	2.913
	1.706		2.652	1.406
KG	7.698	W2G	6.396	0.000
	1.406		0.000	4.603
	1.431		7.490	0.000

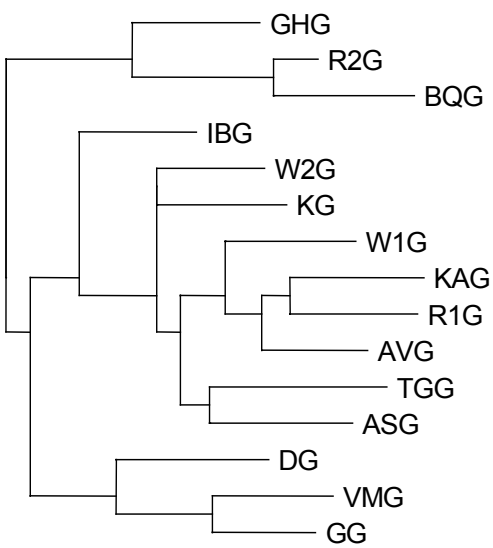

Fig. II-7: Additive trees

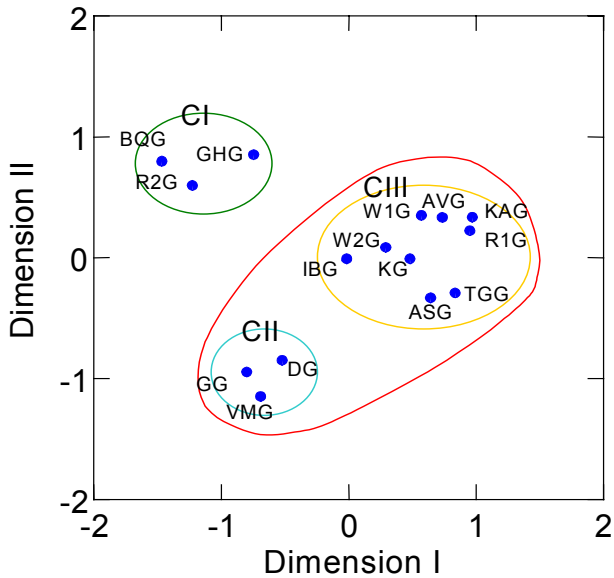

Fig. II-8: Stimulus groupings

The rattling sounds aligned with BQG, R2G, and GHG are heared as high and disrhythmic compared to the accelerating ones aligned with GG, VMG, and DG to be perceived as being loud and high, and to those aligned with ASG, TGG, KG, IBG, W2G, W1G, R1G, KAG, and AVG, all of which are heard as soft and low. The analysis of variance showed that three dimensions explained 79.2% of the variance in the data. The first dimension explained 47.8%, the second dimension 23.5%, and the third dimension described 7.9%. The best contribution to prediction is made by the two first dimensions. Since the third dimension accounts for so little, we may simply ignore it, and concentrate on the plane spanned by the first two dimensions. The interpretation of attributes represented by projections of the stimulus points on principal axes in the coordinate system must be empirically safe. Sensory attributes refer to the properties inferred from verbal or other behavior of observers or ascribed to stimuli by observers. From the verbal reports in questionnaires one finds that Ss rated all sound pairs in terms of such semantic descriptors like loud, quiet, high, dull, rhythmic, accelerating, regular, beating, topic etc., similar to those reported in the previous experiment on preference ratings.

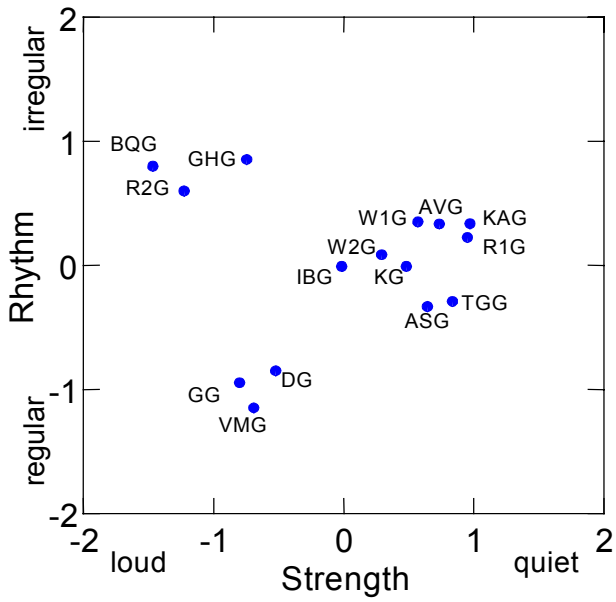

Fig. II-9: A possible solution

The first dimension which might be described by strength is deemed to be a substantial property for pairwise similarity judgments. The second dimension which might be described by rhythm has to do with descriptors of sound attributes like speedy, rhythmic, accelerating, etc. Thus the subjective verdict likely is that the first dimension refers to a ‘loud-soft‘ scale, while the second

dimension refers to a ‘regular-irregular‘ or ‘rhythmic-disrhythmic‘ scale (see Fig. II-9). On the other hand, speculations remain open on whether the third dimension that has to do with descriptive adjectives like high, dull, bright, deep, etc. is related, if any, to tone color.

3.3 Intercorrelations of Perceptual Dimensions

Substantively, we are curious to see whether there are few or much interrelations in the underlying perceptual dimensions used to characterize the sound samples among different types of judgments. To be answered is a question as to how far the locations of stimuli in preference space are predictable from their locations in similarity space, in other words, how far preference ratings depend on the perceptual demensions that govern perceived similarity. As with the hypothesis underlying the unfolding model, the geometric configuration in which the preference relations could be displayed can be thought of as a subspace of a similarity space (e.g. Cooper, 1973). In view of the fact that the Ss give different weights to perceptual dimensions in the psychological spaces of preferences and similarities respectively, with the appropriate statistical analysis of the scaling results, relative ranking of the perceptual dimensions extracted for two sorts of subject’s responses should be comparable. The multiple correlation technique is taken to be appropriate to make an objective interpretation of principal axes in the coordinate systems as sensory attributes or perceptual dimensions (e.g. Green et al., 1969).

Table III-1: Correlations of MDS dimensions of preference ratings with MDS dimensions of similarity judgments

	SIM_DIMI	SIM_DIMII	SIM_DIMIII
PREF_DIMI	0.694**	0.338	–0.273
PREF_DIMII	–0.158	0.289	0.019
PREF_DIMIII	–0.396	0.143	0.125

** correlation coefficient is significant on the level of 0,01 (two-sided).

As for the clues to the common basis of the perceptual dimensions emerging from the similarity judgments with those dominating the preference ratings, we must relate the results of Experiment II to the preference scaling carried out in Experiment I. The comparison of the results from both studies allows us to demonstrate the influence of the nature of sound sources on the perceptual structures that underlie auditory perception. The stimuli’s coordinates on the two orthogonal dimensions of preference ratings and

perceived similarity are respectively given in Table IV-7 in section 4.2. The multivariate correlation technique was implemented to interrelate the results in the preference evaluation to those found in the similarity evaluation in order to see to which degree it accounts for the total variance in the data. Intercorrelations were calculated between metric distances in the geometric representation retained in the similarity data and those in the metric space extracted from preference data. Table III-1 gives multiple correlations between the coordinates on the six perceptual dimensions. In this table it is to note that the correlation of about 70% between the first preference dimension and the first similarity dimension turns out to have statistical significance. Intercorrelations for other cases are negligibly small. The first dimensions pertaining to the similarity judgments and respectively to the preference ratings are related to one another to a certain extent. None of the remaining four correlations reaches statistical significance.

Table III-2: Correlation coefficients between stimuli and NMDS dimensions

Stimulus	PREF_DIMI	PREF_DIMII	PREF_DIMIII	SIM_DIMI	SIM_DIMII	SIM_DIMIII
GG	0.844	0.078	0.263	0.180	-0.840	0.278
AVG	0.955	-0.093	0.018	0.885	0.317	0.051
VMG	0.633	-0.566	-0.282	0.281	-0.796	0.157
ASG	0.838	-0.304	-0.049	0.894	-0.250	0.148
DG	0.903	0.251	-0.272	0.149	-0.880	-0.166
TGG	0.854	-0.197	0.267	0.764	0.076	-0.417
KG	0.894	0.054	-0.262	0.866	0.084	-0.114
BQG	0.787	0.177	0.433	-0.884	0.339	0.231
W1G	0.760	-0.323	-0.202	0.748	0.449	0.059
W2G	0.765	0.242	0.395	0.864	-0.029	0.140
R1G	0.841	0.387	0.012	0.850	0.165	0.075
R2G	0.806	-0.364	0.026	-0.425	0.041	-0.804
KAG	0.896	0.219	0.080	0.830	0.338	-0.020
IBG	0.766	-0.106	-0.125	0.665	-0.127	-0.266
GHG	0.499	0.607	-0.530	-0.013	0.871	0.192

df = 14 *p <.05*

Table III-2 in the Appendix shows the multiple correlations between a set of acoustic sources and perceptual dimensions underlying the multidimensional representations pertaining to the preference and similarity data respectively. Pref_DimI and Sim_DimI showed a positive relationship with regard to those sounds like AVG, ASG, TGG, KG, W1G, W2G, KAG, and IBG, but BQG and R2G, suggesting that the solution retained in the preference study was almost exactly the same as that extracted in the similarity study. Hence, both dimensions are supposed to be associated with the same judgment criterion. The low correlation between Pref_DimII and Sim_DimII may have reflected

different judgment tendencies. The combination of Pref_DimII with Pref_DimIII, of Sim_DimII with Sim_DimIII in each metric space showed no configurational distinction, indicating that the two judgment groups do not differ systematically with regard to the relative importance of the second and third dimension. In reference to Pref_DimII and Sim_DimII, VMG and GHG are only slightly correlated. A comparison of Pref_DimIII with Sim_DimIII shows that in no case the correlation is noteworthy, suggesting that the third dimension is of no substantial predictableness. To compare the results rendered by the multidimensional representations, separate graphs were thus drawn for each combination of generated dimensionality in preference and similarity structures, respectively, and appears in Fig. III.

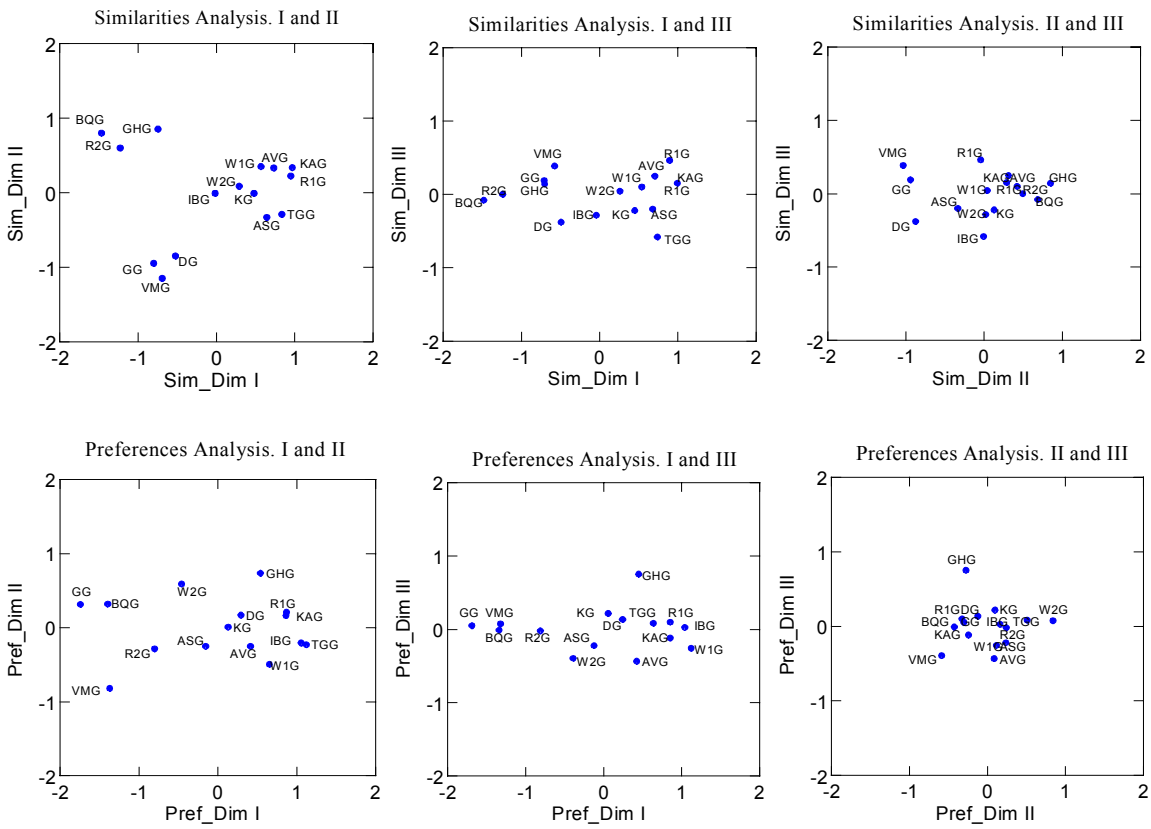

Fig. III: Comparison of scaling spaces based upon similarity and preference Evaluations

Interpretation of Pref_DimI is relatively clear and straightforward. In the final preference structure are those sounds aligned with GHG, R1G, KAG, TGG, IBG, and W1G, all of which are heard as soft, except for DG, in relatively close proximity to one another. Those sounds in the final preference structure like GG, BQG, VMG, and R2G, all of which are perceived as loud, separated by some distance, with the exception of W2G and ASG, whose position is

relatively unstable. Sounds aligned with GG, BQG, VMG, and R2G are rated as loud, whereas those aligned with GHG, R1G, KAG, TGG, IBG, and W1G are heard as more soft and less loud. Pref_DimI would lead us to consider a possible perceptual relation by which a loud-soft scale might be relevant in determining the nature of the final configuration of stimulus points. Sounds aligned with W2G, ASG, DG, KG, and AVG are somewhat neutral concerning Pref_DimI, and its position may be unstable. Including Pref_DimIII shifted all sounds towards the same quadrant, as in the final configuration, without causing major changes in the configuration. Irrespective of GHG, which is a sound similar to a heartbeat, all sounds do not project strongly on the third dimension. Hence Pref_DimIII appears to add no new interpretable information.

In spite of the emergence of Pref_DimI and Pref_DimII, which can be taken as a sign of pronounced sensory attributes, 75.3% of the variance was explained. However, there still are some uncertainties. Pref_DimIII operates obscurely since the explained variance is very low. Pref_DimII does somewhat affect the preferences but the effect is hardly impressive enough to be considered. Except for GHG in the final similarity structure, BQG and R2G are in close proximity to one another, both of which are heard as loud and disrhythmic. Stimuli like BQG, R2G, and GHG are perceived as regular or fluctuating, with a slight edge going to BQG that is heard as purely disrhythmic. GHG, which is discerned as less loud, most likely appears in this quadrant due to its close thematic association with BQG. GG, VMG, and DG are all heard as loud and continuous. Such engine sounds, so to say, are judged to be maximally similar to each other and therefore emerge at opposite ends of what might be roughly interpreted as strength and rhythm. Hence we are led to speculate that loud-soft and regular-irregular scales are the major distinguishing sensory attributes pertaining to the similarity judgments. Although Sim_DimI and Sim_DimII have their own unique characteristics, Sim_DimIII is the minority aspect. Adding Sim_DimIII would not destroy the relatively clear interpretation of the first two dimensions, which, accounting for 71.3% of the variation, were appropriate to represent the psychological space for perceived similarity.

Sensory attributes that give rise to similarity judgments seem to appeal to those pertaining to the preference ratings. This may hint that our subjects assessed more differentially and systematically in pairwise similartiy ratings than has been the case for pairwise preference ratings. The results offered by the data can be considered in the light of the multidimensional unfolding model: The preference structure can be thought of as a subspace of the similarty structure. In terms of similarity evaluation, the hypothesis just mentioned would be supported for reasons that the characteristics used in the preference evaluation can be reflected in the similarity structure. The overall relations seem to be that the first two dimensions pertaining to the similarity judgments were perceived to have high correlations with the perceptual

dimensions pertaining to the preference ratings. In contradiction to the preference case, the factors that promt sound judgments on the extent of similarity appear to impede the judgments on the basis of finer perceptual criteria, by paying attention to specific features of the auditory events. Along with Harrison's concept of response competition (Harrison, 1968), it is suggested that an unfamiliar or novel stimulus tends to give rise to the responses associated with a state of tension and negative affect. As such, by analogy, the subject comes to regard the stimuli as similar or different by ferreting out cues that confirm or disconfirm specific cognitive hypotheses about the nature of stimuli. Relying Coombs (1951), furthermore, it may be stated that the preference ratings are a type of disjunctive behavior with which an individual performs his/her task with the maximal amount of one latent attribute, while the similarity judgments act conjunctively in the sense that two or more latent attributes are required, each to a minimum degree.

3.4 Discussion

We conducted two experiments, and here the Euclidean model appeared to fit quite well: Two dissimilarity matrices were analyzed using Kruskal's procedure, thereby producing two multidimensional configurations. Both solutions achieved satisfactory Stress in 2D Euclidean spaces. It has been mentioned in the introduction that humans just select three or four sensory attributes during the process of auditory perception even though a variety of different perceptual dimensions may exist. We have seen that the judgments of preference and perceived similarity among the stimuli under study may be conceptualized as metric distances in Euclidean space, and this space has an axis system with the stimulus coordinates that are psychologically unique. The choice of principal axes is arbitrary as the model distance does not depend upon this choice. Thus this model is rotationally invariant. Nonmetric scaling analyses of dissimilarity matrices did cause the multidimensional solutions, and thus the resulting pattern of similarities is a bit different than that of preferences. The sounds used in both experiments varied appreciably along the first two dimensions. It should be stressed, however, that one cannot tell from a multidimensional solution for averaged data whether a dimension is moderately important for all subjects or whether it is relevant for some subjects and unimportant for others. From the nonmetric analyses, the first dimension is taken to be perceptually most salient. We find that two dimensions are sufficient to cover much of the variance inherent in the preference and similarity data.

The dimensional interpretation, a matter for speculation, suggests that the psychological distances on the first dimension may be pseudo-physical in its relationship to the perceptive sound level of the stimuli, and those underlying

the second dimension correspond to the perceptive fluctuation. The subjective verdict is that the horizontal dimension could be roughly interpreted as strength and the vertical dimension as rhythm, evident in the similarity data much more than in the preference data, whilst in the preference structure the pronounced sensory attribute strength turns out to be the only perceptual dimension for a reasonable classification of stimuli. As such, the preference ratings may have been dominated only by a loud-soft scale, whilst the perceived similarity depends not only upon strength but also upon rhythm relative to each other. It may be said that the perceived similarity of the same two sounds differing in strength may either be enhanced or decreased by a change in their rhythm. The preference response was tenuously elicited by regularity compared to those found in the similarity evaluation. Faced with a set of sound sources in the similarity evaluation, S seemed to pay attention to more stimulus aspects to define the stimuli than did in the preference evaluation. We are led to speculate that the similarity instruction might have called undue attention to stimulus aspects of the sounds, which otherwise would not have been taken into consideration in the preference ratings. In the present case, the sensory attribute rhythm turned out to be more salient in the similarity evaluation than in the preference ratings. The trend toward separating BQG, R2G, GHG, GG, VMG, and DG from the others in the similarity structure is emphasized much more than in the preference structure as they appear in Fig. III.

It is as though the similarity evaluation was performed more objectively along the first two dimensions than the preference evaluation was. Similarity evaluation thus may be thought of as a 2D surface with strength and rhythm. It may well be that the first preference dimension may relate closely to loud-soft scale, whilst perceived similarity also utilizes this characteristic with an additional perceptual criterion that is temporal in nature. As such, the mental representation of preference ratings and perceptual similarities can be linked to sensory attributes that can be correlated with physical parameters of the sounds. In addition, since the ability to discriminate between the stimuli must necessarily involve the recognition of similarity (e.g. McGeogh & Irion, 1952; Eisler & Ekman, 1959; Gregson, 1975), the perceptual dimensions extracted from nonmetric scaling analysis of perceptual similarities may function as central attributes of sound assessment. It is very likely that the perceived similarity is introspectively based on more distinct characteristics of sound samples, i.e. they require more attention to fine details. It is likely that at higher levels of cognitive processing some categorization process might be effective in the perceived similarity. To summarize the discussion up to this point it may be said that both preference ratings and perceptual similarities have the sensory attribute reflecting strength in common but differ in that the sensory attribute reflecting rhythm becomes more important in influencing the similarity judgments.

4. Instrumental Analysis of Sounds

Even though there is too great a chance that stimulus specific factors will influence the perceptual ratings, both nonmetric multidimensional scaling and nondimensional additive tree show us that the sensory attributes strength and rhythm appear as appropriate to describe the stimuli. At this point, it is worth noting that *"Multidimensional scaling techniques are attractive . . . since the researcher need not physically construct his stimuli to conform to the test of a specific a priori hypothesis. Rather, he may start with the perceptual judgments of similarity among a diverse set of (naturalistic) stimuli, and then explore the various factors which contributed to the subjective distance relationships. These factors may be physical parameters of the stimuli, which then would lead to a psychophysical model; yet, multidimensional scaling techniques may also uncover any other factors involved in judgment strategies." (Grey, 1977, 1270)*. Anyhow, even when appropriately rotated, the coordinate axes do not necessarily provide the most interpretable features of the stimuli. The emergence of the dimensions can possibly be an artifact of the multidimensional scaling operation, suggesting a communality problem stating that each stimulus introduces, if anything, extra and redundant dimensions of its own (e.g. Abelson & Sermat, 1962, 552).

Of great interest is thus the extent to which the principal axes in the coordinate system constitute a relatively stable background for the perceptual dimensions pertaining to the judgments of preference and perceived similarity, i.e., whether the extracted perceptual dimensions are available as a standard for comparison. It has to be emphasized that the Euclidean model does quite well fit the data under study, but provides no clue to the validity of the assumption that the perceptual dimensions are psychologically fundamental. In this respect, Stewart gives a note that *"Even if we assume that a spatial model is an appropriate representation for the psychological structure of a given set of objects, the true importance of dimensions obtained from multidimensional scaling remains to be determined." (Stewart, 1974, 508)*, requring that the generality of a geometric configuration has to be empircally verified. It can be said that treating human auditory perception metrically may be appropriate only when sensory attributes of a sound applied by humans are obvious, thus Attneave (1950) proposed linear models using Least Squares Regression to relate the observed dissimilarity to physical dimensions of the observed stimulus pair (cf. "psychophysical maps", Lazarte & Schönemann, 1991). The prerequisite to capture the number of dimensions for the evaluation of sounds would be to use multidimensional scaling techniques. Sensory attributes involved in the process of auditory perception can mainly be found on the grounds of results of physical measurement, thus

quantified with acoustic or psychoacoustic magnitudes by using the Least Squares Regression. Even though dimensional as well as nondimensional results, which exhaustively showed the sensory attributes pertaining to the judgments of preference and overall similarity, have been based on the coordinates for points with respect to unrotated axes of the solutions, the following acoustic analyses suggest that the interpretation of the perceptual dimensions strength and rhythm should be replaced by two physical parameters, maximum A-weighted SPL and equivalent sharpness.
In light of a global schema of the perceptual study of vehicle sounds ("preference maps", McAdams et al., 1998), it can be said that the objective and instrumental analyses of the stimuli of interest can make it possible to quantify the sensory attributes extracted from the analysis of nonmetric multidimensional scaling. The search for a relationship between the acoustic analysis of sounds and the perceptual analysis of subjective ratings is an issue to be dealt with in the following sections, which, to be sure, allows us to determine the acoustic and/or psychoacoustic magnitudes that are shown to have high correlations with the stimulus coordinates along the perceptual dimensions underlying the psychological spaces. Since the relative importance of the measured physical magnitudes with regard to the perceptual dimensions are not known a priori, thus a multivariate statistical analysis is required to correlate the evaluation ratings with the physical measures. Once the positions of the stimuli studied in regard to their component dimensions of variability are to be determined instrumentally, we can establish a quantitative matching function. And this will substantially aid the interpretation of perceptual dimensions.

4.1 Calculation Models of Auditory Sensations

Ever since the nineties, a lot of objective engineering metrics have been developed, which correlate well with the subjective perception of product sound quality, permitting the prediction of product sound attributes to which humans have easy of access to aurally adequate sound appreciation. Extensively well-tried in empirical tests, metrics allowing the correlation of technically measured sound attributes with information from sound quality perceptions are available on the market and have proven successful tools for sound quality engineering (e.g. Cortex Electronics, HEAD acoustics GmbH, etc.). The HEAD acoustics Binaural Analysis System (BAS, 1996) is an instrumentation tool which, recently by the Oldenburgians, is being used as a standardized procedure for psychoacoustic magnitude calculation (e.g. Kachur, 1998): The psychoacoustic magnitudes implemented in the BAS are based on a generalization of psychoacoustic facts. At the very least such a sound evaluation simulation has the advantage that psychoacoustic

magnitudes are directly connected with the acoustic objects of interest. The BAS makes it possible to break up a complex sound into its aurally adequate psychoacoustic descriptors, such as loudness, pitch, sharpness, roughness, fluctuation strength, tonality, etc., which represent valuable tools for tailoring a suitable sound product. Bodden pointed to the generality of the psychoacoustic indices, which have been developed mostly on the monaural basis of synthetic signals.

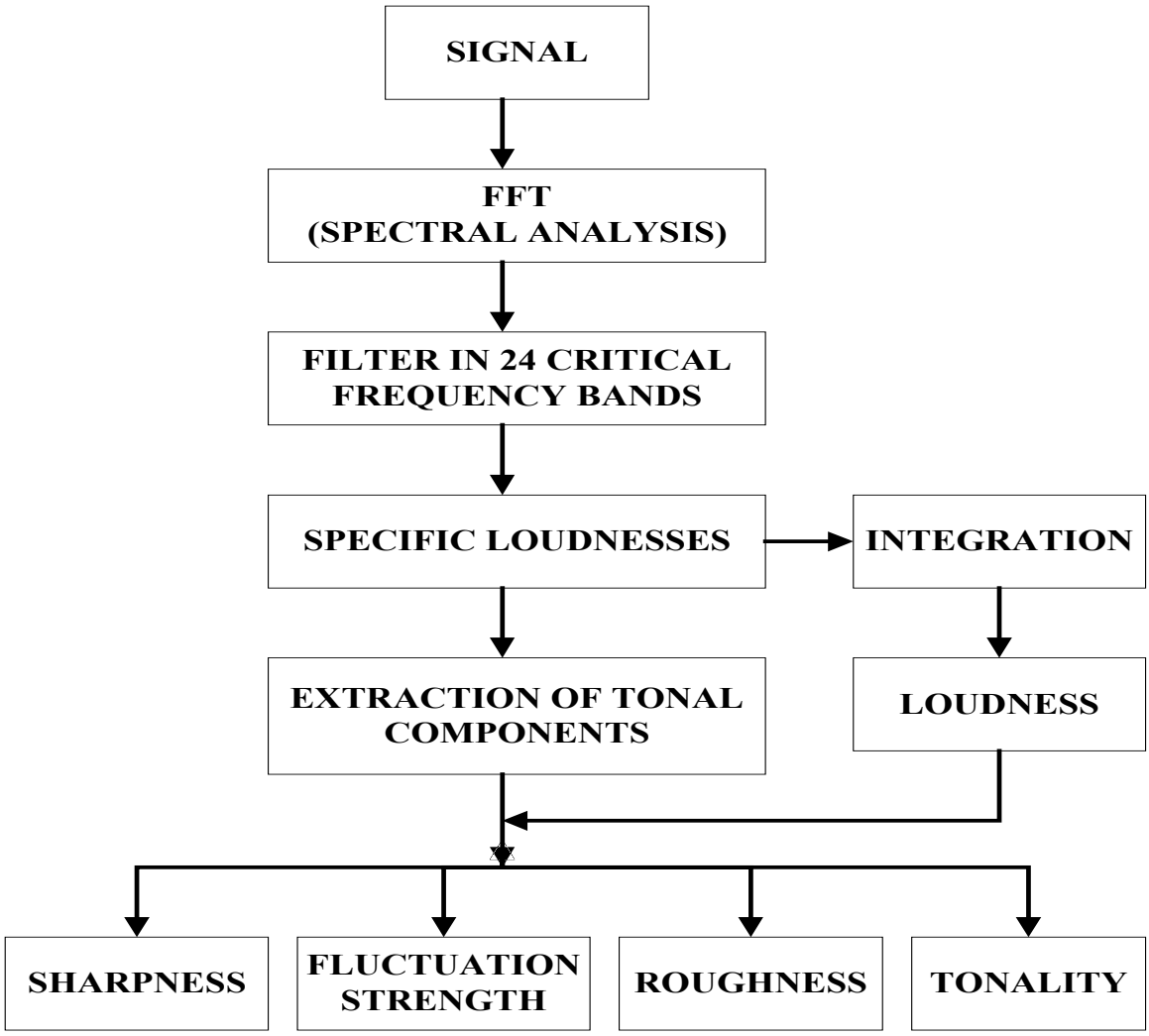

Fig. IV-0: Parameter Evaluation Procedure

Basic hard- and software philosophies of different instrumentation tools for sound quality evaluation including signal analysis, psychoacoustic indices, and signal manipulation have been discussed in the paper by Bodden (1997). Sets of psychoacoustic magnitudes rendered by the BAS are essential thanks to the contribution of Aures (1984a,c, 1985) who showed deviations of 10% between the physical quantities and the psychoacoustic estimates. The HEAD algorithm is substantially based on the calculation model advanced by Aures

who proposed Sensory Pleasantness Index as a combination of the hearing sensations loudness N, roughness R, sharpness S, fluctuation strength F, and tonality K. Later on, Zwicker (1991a,b) advanced the idea of psychoacoustic annoyance, integrating three psychoacoustic parameters loudness, sharpness, and fluctuation strength. As is displayed in parameter evaluation procedure Fig. IV, on the basis of Zwicker's concept of specific loudness, not only loudness but also other time-varying hearing sensations can be derived from temporally variable specific loudness patterns as well. It is not the intention of this section to explicitly consider calculation models proposed for basic auditory sensations that have been presented in the past dozen years and to continue evolving such and. Accordingly, up-to-date developments concerning models of some psychoacoustic parameters which Aures had attempted to integrate to assess Sensory Pleasantness are briefly reviewed. Selected psychoacoustic facts concerning loudness, sharpness, fluctuation strength, roughness, tonality, and sensory pleasantness will be sketched.

4.1.1 Concept of Specific Loudness

To begin with the concept of critical bands proposed by Fletcher who assumed that our hearing system processes sounds in relatively narrow frequency bands that somewhat correspond to the segments on the basilar membrane of the inner ear (cochlea), the frequency spectrum of a tone, by using the third-octave analysis method, is assumed to be partitioned into 24 abutting critical bands which are used for calculating the elementary auditory sensations. At frequencies up to 500 Hz, the critical-band has a constant bandwidth of 100 Hz, and at higher frequencies the critical bandwidth increases a little slower than in proportion to centre frequencies, and at frequencies above 3 kHz a little faster. At frequencies above 500 Hz, a critical bandwidth with approximately 20% of mean frequency corresponds to the relative bandwidth of 23% of a 1/3-octave band filter (Fastl, 1997). The critical-bands piece into 24 filter-banks ranging from 1 Bark to 24 Barks, which corresponds to the audible frequency region ranging from 20 Hz to 16 kHz (Zwicker, 1958). The critical frequency bandwidth Δf_G can be measured as a function of frequency of the tone f:

$$\frac{\Delta f_G}{Hz} = \left[25 + 75 \cdot \left(1 + 1.4 \cdot \left(\frac{f}{kHz} \right)^2 \right)^{0.69} \right] \tag{2.1}$$

Such a relation between the critical bandwidth and the frequency in Eq. (2.1), as a matter of fact, is shown to be be aurally inadequate. In the beginning of the 1980's, Zwicker and Terhardt (1980) proposed the critical-band rate scale

that has the unit "Bark" [22] as an aurally adequate measure for the distribution of specific loudness excitations, and it can be written as follows:

$$\frac{z}{Bark} = \left[13 \cdot \arctan\left(0.76 \frac{f}{kHz} \right) + 3.5 \cdot \arctan\left(\frac{f}{7.5kHz} \right)^2 \right] \tag{2.2}$$

Inversely, the relation between z in Bark and f in Hz can also be expressed as follows (Aures, 1984a, 11):

$$\frac{f}{Hz} = \left(\frac{e^{0.219 \frac{z}{Bark}}}{0.352} + 100 \right) - 32 e^{-0.15 \left(\frac{z}{Bark} - 5 \right)^2} \tag{2.3}$$

4.1.2 Model of Loudness

Having used psychophysical stimuli such as stationary or instationary sounds, a host of authors (e.g. Helmholtz, 1896; Hornbostel, 1926; Kryter, 1967; Weber & Mellert, 1978; Cardozo & Van Lieshout, 1981a,b; Mellert & Weber, 1981; Völk, 1981; Brennecke & Remmers, 1983; Terhardt, 1984; Zwicker, 1960, 1991; Heldman, 1994; Bisping, 1995, 1997; Chouard, 1997; McAdams et al., 1998) reported loudness to be an overriding sensory attribute in influencing subjective responses to sounds. Loudness is the subjective correlate of a sound's "strength". The determination of a loudness function is grounded upon the assumption that several partial loudnesses produced by stimulating the basilar membrane over a wide range corresponding to a length of 1.3 mm on the organ of Corti are summed up to the total loudness in the brain. To calculate loudness, Zwicker's algorithm (Zwicker, 1958, 1960) using pure tones with 94 dB at centre frequencies between 20 Hz and 12.5 kHz in third-octave spacing is generally used: sounds are fed into a FFT (Fast Fourier Transformation)-based measurement system and into a third-octave band analyzer. The envelopes of acoustic signals filtered by critical-band filters are calculated so. The frequency resolution of specific loudnesses along

[22] The Bark was created to honour Barkhausen who had first coined the term "phon" for loudness level. The loudness levels for different frequencies of pure tones are known as equal-loudness contours. The equal-loudness contour for 3 dB at 1 kHz is indicated by L_N = 3 phon (Zwicker & Fastl, 1999, 203). As prescribed from DIN 45 631 (1967), the values of phon can be converted to the values of sone:

$$\frac{L_N}{phon} = \begin{cases} 40 + 33.22 \log \frac{N}{sone}, N \geq 1 sone \\ 40 \left(\frac{N}{sone} + 0.0005 \right)^{0.35}, N < 1 sone \end{cases} \tag{2.4}$$

the Bark scale and the time resolution of the discontinuities of loudness level within 200 ms lead to the total loudness. Over the spectral distribution of the excitation level E(z), specific loudness N′ (loudness within a band) that has the unit "sone/Bark" (in Latin "sonare") depends on critical-band rate:

$$\frac{N'(z)}{sone\,/\,Bark} = N'_0\left(\frac{1}{s}\cdot\frac{E_{HS}(z)}{E_0}\right)^{0.23}\cdot\left[\left(1-s+s\cdot\frac{E(z)}{E_{HS}(z)}\right)^{0.23}-1\right] \tag{3.1}$$

where E_0 stands for a reference value of sound intensity $I_0 = 10^{-16}$W/cm^2, E(z) for the excitation level on the area of z, $E_{HS}(z)$ for the hearing threshold, $s \approx 10^{(0.22-0.005\cdot z/\mathrm{Bark})} - 1$ (e.g. s = 0.5 for frequencies near 1 kHz, Zwicker & Fastl, 1999, 224) is the threshold factor, and $N'_0 \approx 0.068$ sone/Bark as the reference loudness chosen for a sound pressure level (SPL) of 40 dB at 1 kHz. Excitation equals zero, i.e. E = 0, leads to a specific loudness equal to zero, i.e. N′ = 0. Following the Calculation Method of Specific Loudness recommended by ISO 532 (1975) or DIN 45 631 (1967):

$$\frac{N'}{sone\,/\,Bark} = K_1 10^{0.1e_1 L_{HS}}[(0.75+0.25\cdot 10^{0.1(L_E-L_{HS})})^{e_1}-1] \tag{3.2}$$

where L_E is corrected decibel, the exponent e_1 is 0.25, and the constant K_1 is calculated as being 0.0635. L_{HS} is the hearing threshold and can be calculated by the following formula:

$$\frac{L_{HS}}{dB} = 3.64e^{-0.8\ln\frac{f}{kHz}} - 6.5e^{-0.6\left(\frac{f}{kHz}-3.3\right)^2} + 0.001\left(\frac{f}{kHz}\right)^4 \tag{3.3}$$

HEAD algorithm modified Eq. (3.2) with exponent e_2 as 0.23 and constant K_2 as 0.088 and can be written as the following formula:

$$\frac{N'}{sone\,/\,Bark} = K_2\left(\frac{1}{s}\right)^{e_2} 10^{0.1e_2 L_{HS}}[(1-s+s\cdot 10^{0.1(L_E-L_{HS})})^{e_2}-1] \tag{3.4}$$

Integrating the specific loudnesses over all critical-band rates up to 24 Barks, thus total loudness N is the integral within all the specific loudnesses that are produced at different critical-band rates:

$$N = \int_{z=0}^{24\,Bark} N'(z)dz \tag{3.5}$$

The reference loudness is chosen for a 1 kHz tone with a level of 40 dB, which corresponds to 1 sone. Above 40 dB, when enhancing the SPL by 10 dB the sensation of loudness increases by a factor of about two, thus SPL of 40 dB has to be increased to 50 dB in order to double the loudness sensation, which corresponds to 2 sone (Zwicker & Fastl, 1999; Fastl, 1998). While increasing the sound level from 40 dB up to 80 dB, the loudness difference can be just noticeable at a factor of 16. Terhardt (1968a) estimated the just noiticeable loudness difference to be about 1.09, i.e. the change of loudness from 20 down to 10 sone might be perceived by five perceptual steps of loudness difference, about 9% for each step. As a rule of thumb, the perceived loudness of pure tones at 1 kHz is given as a function of the sound pressure level L [23], in its simplest form. It can be written:

$$\frac{N}{sone} = 2^{\frac{(\frac{L}{phon} - 40dB)}{10dB}} \tag{3.6}$$

4.1.3 Model of Sharpness

A series of psychophysical studies confirm strong participation of sharpness in human auditory sensation (e.g. Plomp, 1970; Bismarck, 1972, 1974a,b; Vogel, 1974, 1975; Terhardt, 1968b, 1974, 1981; Terhardt & Stoll, 1981; Zwicker & Terhardt, 1980; Aures, 1984a; Preis, 1987; Kohler-Kotterba [92]; Widmann, 1992; Weber, 1990; Bisping & Giehl, 1996; Fastl, 1997; Daniel, 1997). An annoying metallic coloration, to a large extent, may cause the auditory sensation sharpness. Narrowband sounds having spectral components at high frequencies would be heard sharp or shrill. The factors affecting the perception of sharpness are related to the spectral envelope represented in the specific loudness pattern versus critical-band rate or centre frequency (Zwicker & Fastl, 1999), i.e. sharpness increases approximately linearly with the critical-band rate for centre frequencies from 200 Hz (2 Bark) to about 3 kHz (16 Bark). Referring to the fact that sharpness (in German “Schärfe”) changes almost by a factor of 20 as SPL increases from 30 dB up to 90 dB, the dependence on SPL intervals is considered to be

[23] Since sound is a carrier of mechanical energy, the sound pressure level (SPL) can be expressed by

$$L = 20 \cdot \log\left(\frac{P_{eff}}{P_{ref}}\right) dB = 10 \cdot \log\left(\frac{I_{eff}}{I_{ref}}\right) dB \tag{3.7}$$

where P stands for the measured sound level and P_0 is a reference sound level having the value $2 \cdot 10^{-5}$Pa (Pa = N/m^2) = $2 \cdot 10^{-4}$dyn/cm^2 (ISO/R 131). I is the measured sound intensity and I_0 is the reference sound intensity having the value 10^{-12}W/m (Fricke et al., 1983). For same level L, broadband sounds are louder than narrowband sounds by a factor of three (Fastl, 1998, 4).

relatively unimportant in sharpness. Summed up over the products of critical-band, loudness density, and sharpness exited by narrowband sounds measured at constant loudness, v. Bismarck (1972, 1974a) has undertaken a factorial investigation on descriptive adjectives of timbres of steady sinusoidal tones and proposed a calculation model of relative sharpness that has the unit "acum" (in Latin "acer"). By definition,

$$\frac{S}{acum} = C \cdot \frac{\int_{z=0}^{24\,Bark} N'(z) \cdot g(z)dz}{\int_{z=0}^{24\,Bark} N'(z)dz} \tag{4.1}$$

$$= C \cdot \frac{\int_{z=0}^{24\,Bark} N'(z) \cdot g(z)dz}{N} \tag{4.2}$$

where N′(z) is the loudness density integrated over critical-band rates, and N is referred to a tone at 700 Hz of N = 13 sone corresponding to L = 77 dB. C ≈ 0.11 is a proportionality constant permitting the calibration to the reference sound producing 1 acum, which is a narrowband sound at a centre frequency of 1 kHz having a SPL of 60 dB, with bandwidth $\Delta f < 160$ Hz. The integral in the denominator of Eq. (4.2) represents the total loudness. The numerator of Eq. (4.1) means a weighted first moment of specific loudness in critical-band rate pattern. In order to exclude dependencies of sharpness exited by narrowband sounds upon SPL, the denominator of Eq. (4.2) would be divided through the total loudness. With the reference value of sharpness S_0, $g(z) = 0.111 \cdot S/S_0$, is a weighting function proportional to the sharpness of a narrowband sound at constant loudness. The curve of g(z) can be approached by a following approximation: if z < 15 Bark then g(z) = 1, over 15 Bark g(z) increases exponentially thus $g(z) = 0.2 \cdot e^{0.308(z-15)} + 1$ if $z \geq 15$ Bark. Instead of integrating the product through total loudness N, as is done by v. Bismarck, and later by Zwicker & Fastl (1999), taking the dependencies on specific loudness into account, Aures (1985) has modified the denominator in Eq. (4.1) in order to achieve the variable relations between the specific loudness and the relative sharpness. This can be written as follows:

$$\frac{S}{acum} = C \cdot \frac{\int_{z=0}^{24\,Bark} N'(z) \cdot g'(z)dz}{\ln\left(\frac{N/sone + 20}{20}\right) sone} \tag{4.3}$$

where C is constant for the reference sharpness chosen for 60 dB at 160 Hz, which yields 1 acum. $g'(z) = g(z) \cdot z$ is a weighted moment of the loudness in critical-band rate pattern for the calculated relative sharpness, i.e. specific loudnesses at higher frequencies are much more effective, and reads:

$$g'(z) = 0.0165 e^{0.171 \frac{z}{Bark}} \tag{4.4}$$

To live up the sharpness to sounds with different loudness, the denominator of Eq. (4.2) has been replaced by a loudness-dependent weighting function. HEAD algorithm gives a reformulation of (4.3):

$$\frac{S}{acum} = C \cdot \frac{\int_{z=0}^{24\,Bark} N'(z) \cdot g'(z) dz}{\ln\left(0.05 \frac{N}{sone} + 1\right) sone} \tag{4.5}$$

4.1.4 Model of Fluctuation Strength

Temporal variations of sounds manifest themselves in two auditory sensations: Fluctuation strength at slow variations, and roughness at fast variations (Fastl, 1982, 1983, 1997). Sounds for slow modulation frequencies around 20 Hz, i.e. $f_{\text{mod}} < 20$ Hz, lead to the hearing sensation fluctuation strength (in German "Schwankungsstärke") that has the unit "vacil" (in Latin "vacillare"). A tone at 1 kHz of 60 dB, with degree of modulation m = 1, say 100% amplitude modulated, is referred to as a fluctuation strength of 1 vacil. In sound quality engineering, fluctuation strength is used for sounds with alerting character like warning signals (e.g. see Fastl, 1998). To evoke the perception of maximum fluctuation strength, a degree of modulation of 94%, i.e. m = 0.94, is necessary, which corresponds to an SPL of about 30 dB. No fluctuation strength is perceived up to variations in the temporal envelope of about 3 dB, corresponding to a degree of modulation of 10% (i.e. m = 0.1). The maximum frequency resolution of the human hearing system is about 4 Hz. Maximum fluctuation strength at $f_{\text{mod}} \approx 4$ Hz is like a normal speaking rate, 4 syllables/second. Both fluctuation strength and roughness are calculated using temporal masking patterns, describable as the level difference ΔL that represents the specific loudness ratio between the maximum and the minimum of the level at threshold of the test tone when measured during one period of modulation (Fastl, 1998, 56-57). Fluctuation strength can be described in the equation (Zwicker & Fastl, 1999, 256):

$$\frac{F}{vacil} = \frac{0.008 \int_{z=0}^{24\,Bark} \left(\frac{\Delta L}{dB} Bark \right) dz}{\left(\frac{f_{\text{mod}}}{4Hz} \right) + \left(\frac{4Hz}{f_{\text{mod}}} \right)} \tag{5.1}$$

Instead of the integral in Eq. (5.1), it can be replaced by a sum of 240 partial level differences between maximal and minimal specific loudness values in each of the 21 channels of the loudness meter (Zwicker et al., 1985) along the critical-band rate scale. Thus HEAD reformulated Eq. (5.1) as follows:

$$\frac{F}{vacil} = \frac{\frac{0.36}{Bark} \cdot \int_{z=0}^{24\,Bark} \log\left(\frac{N'_{\max}}{N'_{\min}} \right) dz}{\left(\frac{T}{0.25s} \right) + \left(\frac{0.25s}{T} \right)} \tag{5.2}$$

where N'_{max} and N'_{min} are maximal and minimal specific loudness values in each channel of a loudness meter, and T described as "s" refers to the temporal difference between the two successive loudness maxima.

4.1.5 Model of Roughness

For fast modulation frequencies in the region between about 15 Hz to 300 Hz, i.e. 15 Hz < f_{mod} < 300 Hz, fluctuation strength vanishes and roughness takes over (e.g. Terhardt, 1968b; Vogel, 1974; Aures, 1984b; Daniel & Weber, 1993; Daniel, 1997; Widmann & Fastl, 1998). The minimum roughness of 0.1 asper is reached when the degree of modulation amounts to about 25%, i.e. m = 0.25. An increment in the degree of modulation of 10% corresponds to an increment in roughness of 17%, thus there are about 20 audible roughness steps. According to Terhardt (1968), the just noticeable roughness difference refers to about a factor of 1.56. Roughness (in German "Rauhigkeit") that has the unit "asper" (in Latin "vox aspera") is related to the degree of modulation, i.e. $R \supseteq m^p$, $1.5 \leq p \leq 2$. Roughness increases almost in proportion to the square of the degree of modulation. In line with the model of sensory pleasantness, Aures (1984a,c) proposed a roughness calculation procedure, under the assumption that total roughness is like the integral within 24 specific roughnesses:

$$\frac{R}{asper} = C \cdot \int_{z=0}^{24\,Bark} r'(z)dz \tag{6.1}$$

where the constant C is used to standardize the calculated roughness values on the reference sound. This is a 100% amplitude modulated 1 kHz tone with a level of 60 dB and a modulation frequency of 70 Hz, which elicits a roughness of 1 asper. r′(z) refers to specific roughnesses along the critical-band rate scale, and this can be described by the equation:

$$r'(z) = C \cdot m^2 \frac{asper}{Bark} \tag{6.2}$$

Roughness can be expressed by the product of modulation frequency and level difference, i.e. $R \approx f_{\mathrm{mod}} \cdot \Delta L$. The product reaches a maximum for a modulation frequency near 70 Hz. Above 70 Hz, ΔL is decreasing rapidly and therefore roughness disappears. The total roughness R is calculated from the excitation level difference and the modulation frequency (Zwicker & Fastl, 1999, 263), which was implemented in HEAD algorithm:

$$\frac{R}{asper} = 0.3 \frac{f_{\mathrm{mod}}}{kHz} \int_{z=0}^{24\,Bark} \frac{\Delta L_E(z)dz}{dB\,/\,Bark} \tag{6.3}$$

By a level difference of 40 dB both fluctuation strength and roughness increase by a factor of three. Vogel (1974) set the level difference to be maximal 30 dB for the roughness calculation procedure.

4.1.6 Model of Tonality (Tonalness)

Distinguishing sounds from noises [24] is attributable to the fact that the tonal components play an important role, and this leads to the evolvement of a psychoacoustic parameter termed tonality (in German "Tonhaltigkeit") that has the unit "tu" (abbreviated from tonality unit). Tonality is like a feature distinguishing noise versus tone quality of sounds. Relative tonality depends

[24] *"Noise is sound occuring within the frequency range of human hearing which disturbs silence or an intended sound perception and results in annoyance or endangers the health" (DIN 1320).* The distinction between sound (in German "Klang") and noise (in German "Geräusch") is popularly credited to Helmholtz (1863) who took the view that such primary perceptual categories could be accounted for by those physical components like amplitude, pitch, and timbre. Commonly, "noise" is understood as unwanted sound (in German "Lärm") or as masking sound (in German "Rauschen"). As a matter of fact, complex noise problem is a purely multidimensional phenomenon. From a logographical analysis of letters of noise complaint, for instance, Laucken, Mees, and Schick (1987) came to the conclusion that complex noise problems are mostly social problems.

on the bandwidth expressed in critical-band rate spread (Zwicker & Fastl, 1999). Maximum tonality is achieved at about 700 Hz. A sound having spectral components is called tonal especially when its bandwidth is smaller than the width of a frequency group. The auditory sensation of tonality would be extant just when the prominence ratio measure is larger than 6 dB, i.e. the level of tonal frequency group measured on the 1/3-octave analysis is 6 dB larger than the arithmetical mean of the level measured from both neighboring frequency groups (DIN 45 681, 1992). As such, the gain of tonality amounts to 10 dB. Aures (1985) applied the algorithm for capturing the holistic pich percept, say virtual pitch (measured in pitch units, 'pu') proposed by Terhardt (1979) and Terhardt, Stoll, & Seewann (1982), to the calculation of tonality, the procedure for which can be approximated:

$$\frac{K}{tu} = \sqrt{\sum_{i=1}^{N} \left[W_1(\Delta z_i) W_2(f_i) W_3(\Delta L_i) \right]^2} \tag{7.1}$$

where the term $W_1(\Delta z_i)$, which represents the dependence on differences of critical bandwidth of i-th tonal components, is expressed as:

$$W_1(\Delta z_i) = \left(\frac{0.13}{\Delta z / Bark + 0.13} \right)^{1/0.29} \tag{7.2}$$

Tonality decreases with increasing critical bandwidth, and below 500 Hz tonality begins to decrease. The term $W_2(f_i)$, which describes the dependence on frequency of i-th tonal components, can be approximated as follows:

$$W_2(f_i) = \frac{1}{\sqrt{1 + 0.2(f_i / 0.7kHz + 0.7kHz / f_i)^2}} \tag{7.3}$$

Finally, the term $W_3(\Delta L_i)$ estimates the effect of level surplus of i-th tonal components, and reads:

$$W_3(\Delta L_i) = \left(1 - e^{-\frac{\Delta L_i}{15dB}} \right) \tag{7.4}$$

The rise in the absolute level surplus from 20 dB up to 30 dB has little influence on tonality. Bearing in mind that loudness from 14 sone up takes no noticeable effect on tonality, thus tonality of sounds being compared is related to a tone at 1 kHz of N = 14 sone or L = 78 dB. HEAD algorithm reformulates the Eq. (7.1):

$$\frac{K}{tu} = C \cdot W_N^{0.79} \sqrt{\sum_{i=0}^{N-1} [W_1(\Delta z_i) W_2(f_i) W_3(\Delta L_i)]^2}^{\,0.29} \tag{7.5}$$

where the constant C is normalized to a sinusoidal tone at 1 kHz of N = 4 sone or L = 60 dB results in 1 tu. The exponent 0.29 in Eq. (7.5) is empirically optimized which to represent the dependence of tonality on bandwidth and frequency of sounds using one tonal component. W_N describes the relationship between the loudness without tonal components N_N and the loudness having tonal components N, and by definition,

$$W_N = 1 - \frac{N_N}{N} \tag{7.6}$$

4.1.7 Model of Sensory Pleasantness

The perception of consonance vs. dissonance for aesthetic quality of auditory events that has been in discussion since Helmholtz's time (1863) is considered to be a derivative that traces back to several auditory sensations (e.g. Terhardt & Stoll, 1981; Terhardt, 1984; Weber, 1990; Ellermeier et al., 1997) rather than to a function of the critical bandwidth (e.g. Plomp & Levelt, 1965). Extending this idea to study the psychoacoustic factors affecting the perception of sensory euphony, Aures (1984a,c, 1985) developed a metric for sensory euphony (in German "sensorischer Wohlklang") on which those hearing sensations, such as sharpness, roughness, loudness, and tonality, have an impact. Sensory euphony that dates back to Stumpf (1910) who coined the terms euphony (in German "Wohlklang") vs. disphony (in German "Übelklang"), is generally said to decrease with increasing roughness, sharpness, and loudness, whilst it is enhanced by increasing tonality. A relative value of sensory pleasantness for any sounds can be approximated as a function of relative values of roughness R, sharpness S, tonality K, and loudness N against their reference values:

$$\frac{W}{W_0} = e^{-0.7\frac{R}{R_0}} \cdot e^{-1.08\frac{S}{S_0}} \cdot \left(1.24 - e^{-2.43\frac{K}{K_0}}\right) \cdot e^{-\left(0.023\frac{N}{N_0}\right)^2} \tag{8.1}$$

Just from 14 sone and up, sensory pleasantness is not strongly affected by loudness. The reference value of loudness is normalized to N_0 = 1 sone, which

corresponds to a sinusoidal tone at 1 kHz of L = 40 dB. Similarly, R_0 = 1.3 asper for N = 14 sone at 2 kHz of f_{mod} = 70 Hz and m = 1, S_0 = 9.6 acum at 8 kHz of N = 14 sone, and K_0 = 1 tu at 500 Hz of N = 14 sone. The reference value of W_0 is sensory pleasantness of a sinusoidal tone at 500 Hz. With reference to a tone at 1 kHz of L = 60 dB, which yields W = 1, the absolute value of sensory pleasantness on the basis of the calculations of separate values of auditory sensations is as follows:

$$W = e^{-0.55\frac{R}{asper}} \cdot e^{-0.113\frac{S}{acum}} \cdot \left(1.24 - e^{-2.2K}\right) \cdot e^{-\left(0.023\frac{N}{sone}\right)^2} \tag{8.2}$$

On the other hand, Zwicker (1991a,b) and Widmann (1992) proposed a procedure for assessing psychoacoustic annoyance called unbiased annoyance (in German "unbeeinflusste Lästigkeit") that has the unit "au" (abbreviated from annoyance unit). In the unbiased annoyance model (UBA), the tonality metric was excluded from consideration because tonal signals show higher loudness than noisy sounds do for the same levels:

$$\frac{UBA}{au} = d \cdot \left(\frac{N_{10}}{sone}\right)^{1.3} \cdot \left[1 + 0.25 \cdot \left(\frac{S}{acum} - 1\right) \cdot \log\left(\frac{N_{10}}{sone} + 10\right) + \left(\frac{0.3F}{vacil} \cdot \frac{1 + N_{10}/sone}{0.3 + N_{10}/sone}\right)\right] \tag{8.3}$$

where N_{10} is 10% percentiles [25] indicating that specific loudness reaches 10% of the time period, and the day factor, d (≈ 15 dB or 2 sone), weights the night situation and can be approximated as follows:

$$d = 1 + \left(\frac{N_{10}}{5sone}\right)^{0.5} \tag{8.4}$$

Later on, to approximate Zwicker's unbiased annoyance model using five metrics except the day factor, Laux and Davies (1998) have implemented the UBA model in an artificial neural network model (ANN) of noise annoyance. The network performance was measured in mean squared error (MSE < 0.05). From the sub-network training for the 3-input ANN model, with inputs percentile loudness, sharpness and fluctuation strength, it was found that

[25] Such a statistical measure was brought about by the deficiency of L_{eq} (ISO R 1996 as well DIN 45 645)

$$L_{eq} = 20 \cdot \log\left(\int_{t_1}^{t_2} \frac{P_{eff}}{P_{ref}} dt\right) dB \tag{8.5}$$

Weighted levels are not sufficient to evaluate the audible changes in temporal variations in each specific loudness channel (e.g. Widmann, 1992, 1997). The infra sound problem known as the sonic boom in turbojet aircraft or rumble noise in cars occurs in the frequency range of 60-120 Hz (e.g. Van der Auweraer et al., 1997): A negative correlation between weighted levels and listener annoyance was found. Discussions on the concept of L_{eq} are extensively given by Schick (1990b, 1992, 1994, 1995).

linear combinations of sharpness and fluctuation strength did not substantially improve the predictability of noise annoyance over percentile loudness alone (correlation coefficient of 0.9183). The 4-input ANN model, with inputs N_{10}, sharpness, fluctuation strength, and roughness, had a correlation coefficient of 0.9718, indicating that adding the roughness metric improved the performance of the 3-input ANN model significantly.

Taking account of the level fluctuations, various sound levels such as L_1, L_{10}, L_{50}, L_{90} as well L_{95} have been created, indicating that the level exceeds 1%, 10%, 50%, 90% as well 95% of the time period, respectively. To clarify the matter, let us make an example: we have a sound duration of 10 seconds, one second of 80 dB, two seconds of 70 dB, and the remaining seven seconds have 50 dB. Thus it can be expressed in terms of percentiles of loudness: N_{10} = 80 dB, N_{30} = 70 dB, and N_{100} = 50 dB. Regarding the traffic noise, Fastl (1991) considered N_4 to be appropriate, however Yu (1987) gave evidence for N_{50} showing up better in assessing the metropolitan traffic noises (Schick, 1990b, 104). Zwicker defined the term "unbiased annoyance" as *"the response of subjects being annoyed exclusively by sound under describable acoustical circumstances in laboratory conditions without relation to the nature of the sound source." (Zwicker, 1991, 91)*. The adverb "unbiased" thereby is based on the proviso that the source of sounds to be estimated subjectively remains unrecognizable in the course of presentation, i.e. content free. It is accepted practice to avoid the problems of nonacoustic factors like amenity and or recognizability of acoustic sources in psychoacoustic research by presenting a limited set of reasonably similar sounds. However, it it our conviction that sound evaluation studies should deal with sounds known to the subjects. Annoyance is likely to be susceptible to numerous biasing influences.

4.2 Multivariate Regression Analysis

We may now proceed to report statistical analyses of 15 sound samples which have been classified with respect to a set of 20 acoustic and psychoacoustic magnitudes. Analyses were performed using the HEAD acoustics Binaural Analysis System (BAS, 1996). Sound sources were converted into digital signals, which were then stored on hard disk of a computer. Two major aspects of the physical classification of 15 sound samples were considered: First, we looked for physical parameters that were well correlated with the perceptual dimensions pertaining to the preference structure as well the similarity one. Secondly, we simultaneously sought to capture the underlying best-fitting structure in such a way that metric distances on two or three perceptual dimensions combine into those on the basis of two or three acoustic and/or psychoacoustic magnitudes. The percentage BAS measures

were calculated by multiple correlation techniques. In reference to the multiple correlations between any two physical parameters the different combinations of the percentage BAS measures were then analyzed by the Kruskal's procedure. To do this, metric distances for two or three acoustic and/or psychoacoustic magnitudes in different combinations were estimated.

Table IV-1: Instrumental analysis of 15 stimuli using the BAS algorithm (HEAD acoustics GmbH, 1996)

	GG	AVG	VMG	ASG	DG	TGG	KG	BQG	W1G	W2G	R1G	R2G	KAG	IBG	GHG
dB MAX	86	90,7	86,90	86,59	91,7	91,7	89,90	82	90,09	88,2	89,09	86,7	91,3	92	84,59
dB LEQ	82,5	85,2	84,2	82,3	86,59	85,90	88,09	76,40	84,7	84	84,59	83,2	83,8	87,09	71,09
dB(A) MAX	81	67,3	83,2	72	75,40	65,7	72,5	78,09	66,2	75,8	66,7	81,8	61,4	58,8	72,8
dB(A) LEQ	78,5	65,09	80,90	69	70,8	58,5	70,59	71,59	61,9	73	60,5	75	58,9	56,7	55,1
SONE MAX	55	22,8	60,7	27,9	35,6	21	27,5	42,8	20,2	38,3	18	44,3	12,9	13,5	24,6
SONE LEQ	43,7	19,2	50,9	21,8	27,2	12	24,3	20,7	13,7	29,2	11,8	31	11,2	11,7	4,1
VACIL MAX	3,11	2,35	1,91	2,26	1,97	2,37	1,87	4,92	2,47	3,27	2,32	4,18	2,02	2,16	2,51
VACIL LEQ	2,42	2,06	1,71	1,85	1,65	1,93	1,66	3,81	2,04	2,64	1,53	3,57	1,68	1,82	1,59
ACUM MAX	3,61	2,7	4,14	1,47	2,21	1,56	1,17	4,48	1,41	1,91	1,34	3,51	1,34	1,4	4,74
ACUM LEQ	3,16	1,56	3,82	1,3	1,96	1,25	1,07	2,75	1,11	1,64	0,98	2,32	1,2	1,31	1,91
TU MAX	0,245	0,209	0,321	0,194	0,438	0,457	0,5	0,225	0,163	0,261	0,448	0,125	0,2	0,122	0,389
TU LEQ	0,115	0,046	0,206	0,043	0,186	0,101	0,37	0,039	0,023	0,057	38	0,03	0,092	0,064	0,118
ASPERMAX	4,09	1,97	4,29	2	3,94	1,03	1,52	12,2	1,55	3,77	0,786	5,72	0,845	0,767	2,06
ASPERLEQ	3,1	1,62	3,57	1,62	2,78	0,69	1,16	7,4	0,92	2,3	0,516	3,79	0,615	0,633	0,3
IU MAX	8,619	4,9	4,66	8,609	6,77	5,66	0,527	9,630	29,1	6,93	0,9	4,91	0,974	2,67	2,28
IU LEQ	0,35	0,33	0,27	0,34	0,89	0,29	0,13	3,61	0,4	0,41	0,31	0,79	0,231	0,58	0,61
MOD MAX	113,6	150,9	180,4	98,90	331,6	99,2	95,3	181,4	299,8	280,9	132,4	150	106,4	108,3	150,2
MOD LEQ	64	62,3	66,59	64,2	82	61,8	63,4	99	68,5	64,7	68,5	75,2	67,5	72,5	36,2
PRM MAX	13,7	10,2	16,7	10,5	13,2	14,3	10	26,7	9,6	12,5	9,69	17,2	14,2	15,2	15,7
PRM LEQ	7,3	5,2	8,300	5,5	6,8	6,5	6,2	7,8	5,3	7,2	5	5,9	8,300	9,4	10,5

Table IV-2: Description of the selected parameters using the BAS algorithm (HEAD acoustics GmbH, 1996)

unit of quantities	**description/background settings**
dB Max	maximum sound level transformation: FFT FFT size: 4096 weighting function: Hanning time weighting: fast frequency scale: logarithmic
dB L_{eq}	equivalent sound level transformation: FFT FFT size: 4096 weighting function: Hanning time weighting: fast frequency scale: logarithmic

continued from Table IV-2

unit of quantities	description/background settings
dB(A) Max	A-weighted maximum sound level* transformation: FFT FFT size: 4096 weighting function: Hanning time weighting: fast frequency scale: logarithmic dB weighting: A
dB(A) L_{eq}	A-weighted equivalent sound level transformation: FFT FFT size: 4096 weighting function: Hanning time weighting: fast frequency scale: logarithmic dB weighting: A
Sone Max**	maximum loudness calculation method: FFT / ISO 532 FFT size: 4096 weighting function: Hanning sound field: free frequency resolution: 0.5 Bark (48 bands) filter: 4th order
Sone L_{eq}**	equivalent loudness calculation method: FFT / ISO 532 FFT size: 4096 weighting function: Hanning sound field: free frequency resolution: 0.5 Bark (48 bands) filter: 4th order
Vacil Max	maximum fluctuation strength rpm resolution: adaptive high
Vacil L_{eq}	equivalent fluctuation strength rpm resolution: adaptive high
Acum Max	maximum sharpness calculation method: FFT / ISO 532 FFT size: 4096 weighting function: Hanning sound field: free frequency resolution: 0.5 Bark (48 bands) filter: 4th order method: Aures
Acum L_{eq}	equivalent sharpness calculation method: FFT / ISO 532 FFT size: 4096 weighting function: Hanning sound field: free frequency resolution: 0.5 Bark (48 bands) filter: 4th order, method: Aures

continued from Table IV-2

unit of quantities	description/background settings
Tu Max	maximum tonality rpm resolution: adaptive high
Tu L_{eq}	equivalent tonality rpm resolution: adaptive high
Asper Max	maximum roughness calculation method: modulation rpm resolution: adaptive high
Asper L_{eq}	equivalent roughness calculation method: modulation rpm resolution: adaptive high
Iu Max	maximum impulsiveness rpm resolution: adaptive low
Iu L_{eq}	equivalent impulsiveness rpm resolution: adaptive low
Mod Max	maximum modulation degree FFT size: 256 band resolution: third octave selected band: 15 900-1120 Hz maximum envelope frequency: 500 Hz
Mod L_{eq}	equivalent modulation degree FFT size: 256 band resolution: third octave selected band: 15 900-1120 Hz maximum envelope frequency: 500 Hz
Prm Max***	maximum prominence ratio rpm resolution: adaptive low
Prm L_{eq}	equivalent prominence ratio rpm resolution: adaptive low

* Frequency weightings with the hearing threshold (A) or isophones (B, C, D) have been used to change the sound level into a much more aurally adequate measure. Sounds with the same dB(A) value often do not have the same loudness value. Temporal masking that occurs when the perception of sound energy is suppressed by sound energy that precedes it can have a large influence on the discrepancy between dB(A) and loudness value.
** Specific loudness with percentiles N_4, and N_{10} was not included in scaling the sounds, since it was considered that maximums and medians for loudness would be sufficient to settle the point whether or not variations in decibel would have any appreciable weight in determining perceptual ratings.
*** Prominence ratio is seen to be linked to judgments on a scale of defensive to offensive which has to do with tonality or impulsiveness.

The Euclidean metric was used as a dissimilarity measure. Owing to the ordinal transformation, the BAS values may themselves be considered psychological distances. For the sake of the nature of the multidimensional properties, the compatibility of the subjective positions of 15 sound samples with their objective ones underlying the acoustic and/or psychoacoustic magnitudes will be dealt with. Consider a symmetric matrix that consists of 20 different physical parameters relating the 15 sound samples (see Table IV-

1 as well as Table IV-2). A spatial representation was sought for a set of the BAS values consisting of 20 variables. If the statistical distributions inherent in the data are not known beforehand, it would be legitimate to normalize the raw values in reference to the largest values which occur in each parameter of analysis. For the purpose of standardizing the BAS measures, each is based on a distinct unit of measurement. The data had been converted by $x_i/max(x_i)$ *% 100* tranformations. Thus the value ranges from 0 to 100 so as to indicate the extent to which each physical parameter would describe a specific acoustic and/or psychoacoustic property of human hearing relating to the 15 sound samples. The resulting data matrix contains 380 cells with a column for each physical parameter and a row for each sample in interest (see Table IV-3). We then have carried out the two-sided multiple correlation analysis to obtain the correlations for the 380 pairs of 20 acoustic and/or psychoacoustic magnitudes (see Table IV-4).

Table IV-3: Standardization in reference to the largest values which to occur in each observed parameter

(x_i/max(x_i)*100)

	GG	AVG	VMG	ASG	DG	TGG	KG	BQG	W1G	W2G	R1G	R2G	KAG	IBG	GHG
dB MAX	93,48	98,59	94,45	94,13	99,67	99,67	97,72	89,13	97,93	95,87	96,84	94,24	99,24	100	91,95
dB LEQ	93,64	96,70	95,56	93,42	98,3	97,5	100	86,72	96,14	95,34	96,03	94,44	95,12	98,86	80,7
dB(A) MAX	97,36	80,89	100	86,54	90,63	78,97	87,14	93,87	79,56	91,11	80,17	98,31	73,8	70,67	87,5
dB(A) LEQ	97,03	80,47	100	85,29	87,52	72,31	87,27	88,5	76,51	90,23	74,78	92,70	72,81	70,09	68,11
SONE MAX	90,61	37,56	100	45,96	58,65	34,6	45,3	70,51	33,28	63,1	29,65	72,98	21,25	22,24	40,53
SONE LEQ	85,84	37,72	100	42,83	53,44	23,58	47,74	40,67	26,92	57,37	23,18	60,9	22	22,99	8,060
VACIL MAX	63,21	47,76	38,82	45,93	40,04	48,17	38,01	100	50,2	66,45	47,15	84,95	41,06	43,9	51,02
VACIL LEQ	63,52	54,07	44,88	48,56	43,31	50,66	43,57	100	53,54	69,29	40,16	93,7	44,09	47,77	41,73
ACUM MAX	76,16	56,96	87,34	31,01	46,62	32,91	24,68	94,51	29,75	40,3	28,27	74,05	28,27	29,54	100
ACUM LEQ	82,72	40,84	100	34,03	51,31	32,72	28,01	71,99	29,06	42,93	25,65	60,73	31,41	34,29	50
TU MAX	49	41,8	64,2	38,8	87,59	91,40	100	45	32,6	52,2	89,59	25	40	24,4	77,8
TU LEQ	0,3	0,12	0,54	0,11	0,49	0,27	0,97	0,1	0,06	0,15	100	0,08	0,24	0,17	0,31
ASPERMAX	33,52	16,15	35,16	16,39	32,3	8,44	12,46	100	12,7	30,9	6,44	46,89	6,93	6,29	16,89
ASPERLEQ	41,89	21,89	48,24	21,89	37,57	9,32	15,68	100	12,43	31,08	6,97	51,22	8,310	8,550	4,05
IU MAX	29,62	16,84	16,01	29,59	23,26	19,45	1,81	33,09	100	23,81	3,09	16,87	3,35	9,18	7,84
IU LEQ	9,69	9,140	7,48	9,42	24,65	8,029	3,6	100	11,08	11,36	8,59	21,88	6,4	16,07	16,9
MOD MAX	34,26	45,51	54,4	29,83	100	29,92	28,74	54,7	90,41	84,70	39,93	45,24	32,09	32,66	45,3
MOD LEQ	64,65	62,93	67,27	64,84	82,83	62,42	64,04	100	69,19	65,34	69,19	75,95	68,18	73,23	36,57
PRM MAX	51,31	38,2	62,55	39,33	49,44	53,56	37,45	100	35,96	46,82	36,33	64,42	53,18	56,93	58,8
PRM LEQ	69,52	49,52	79,05	52,38	64,76	61,9	59,05	74,29	50,48	68,56	47,62	56,19	79,05	89,52	100

One can plot the perceptual dimensions revealed in preference as well as similarity studies tested against the acoustic and/or psychoacoustic magnitudes relating the 15 sound samples. Of course, we had no way of knowing in adavance which acoustic and/or psychoacoustic magnitudes, or which combination of acoustic and/or psychoacoustic magnitudes whatsoever, are best correlated with the perceptual dimensions extracted from the perceptual structures. With regard to the problem of relating sensory attributes to component dimensions, the multiple correlation technique might be useful for determining the signal parameters that are correlated with the position of

each sound source along each perceptual dimension, i.e. to discover the common stimulus space which can best account for the perceptual spaces obtained by using the rating procedures.

Table IV-4: Multiple correlations among acoustic and psychoacoustic parameters

	DB MAX	DB LEQ	DBA MAX	DBA LEQ	SONE MAX	SONE LEQ	VACIL MAX	VACIL LEQ	ACUM MAX	ACUM LEQ	TU MAX	TU LEQ	ASPER MAX	ASPER LEQ	IU MAX	IU LEQ	MOD MAX	MOD LEQ	PRM MAX	PRM LEQ
DB MAX	1																			
DB LEQ	0,791**	1																		
DBA MAX	-0,676**	-0,287	1																	
DBA LEQ	-0,41	0,128	0,871**	1																
SONE MAX	-0,586*	-0,158	0,946**	0,917**	1															
SONE LEQ	-0,271	0,216	0,794**	0,951**	0,91**	1														
VACIL MAX	-0,687**	-0,459	0,473	0,328	0,41	0,136	1													
VACIL LEQ	-0,578*	-0,279	0,491	0,436	0,456*	0,252	0,966**	1												
ACUM MAX	-0,778**	-0,733**	0,69**	0,381	0,664**	0,36	0,527*	0,456	1											
ACUM LEQ	-0,595*	-0,31	0,806**	0,73**	0,914**	0,785**	0,37	0,39	0,815**	1										
TU MAX	0,143	0,066	0,067	-0,076	-0,032	-0,064	-0,388	-0,474	-0,091	-0,124	1									
TU LEQ	0,058	0,085	-0,192	-0,22	-0,244	-0,221	-0,107	-0,24	-0,243	-0,275	0,354	1								
ASPER MAX	-0,728**	-0,425	0,655**	0,555*	0,635*	0,379	0,837**	0,83**	0,664**	0,636*	-0,204	-0,219	1							
ASPER LEQ	-0,657**	-0,268	0,699**	0,684**	0,721**	0,535*	0,767**	0,8**	0,615*	0,703*	-0,229	-0,23	0,976**	1						
IU MAX	-0,061	-0,003	0,026	0,066	0,053	0,018	0,181	0,209	-0,057	-0,01	-0,367	-0,227	0,137	0,131	1					
IU LEQ	-0,6**	-0,475	0,276	0,152	0,246	-0,037	0,761**	0,704**	0,493	0,331	-0,171	-0,109	0,89**	0,811**	0,138	1				
MOD MAX	0,062	0,041	0,23	0,21	0,199	0,152	0,098	0,115	0,055	0,104	-0,011	-0,118	0,242	0,223	0,561*	0,173	1			
MOD LEQ	-0,09	0,244	0,163	0,361	0,233	0,229	0,498	0,576*	-0,01	0,205	-0,253	0,014	0,664**	0,72**	0,193	0,692**	0,275	1		
PRM MAX	-0,616*	-0,526*	0,381	0,211	0,416	0,132	0,707**	0,674**	0,681**	0,586*	-0,21	-0,272	0,836**	0,777**	-0,089	0,846**	-0,022	0,51	1	
PRM LEQ	-0,239	-0,51	0,003	-0,198	0,075	-0,085	-0,003	-0,084	0,462	0,335	-0,039	-0,343	0,109	0,023	-0,325	0,178	-0,108	-0,236	0,508	1

** correlation coefficient is significant on the level of 0,01 (two-tailed).
* correlation coefficient is significant on the level of 0,05 (two-tailed).

Once a certain parameter is shown to be highly correlated with another, the interrelated parameters were taken into the nonmetric scaling analysis using Kruskal's procedure, as had been explained in the foregoing sections. Finally, the degree of preference ratings and perceived similarity associated with each stimulus as a function of the perceptually significant parameters was estimated. In doing so, the spatial representation provides a tool by which one can improve the sound space of vehicle interiors according to criteria associated with auditory quality (e.g. McAdams et al., 1998; McAdams et al., 1999; Susini et al., 1999). Since the perceptual spaces had two or three common sensory attributes, we tested different pairs of acoustic and/or psychoacoustic magnitudes for 2D preference as well as similarity space. To start with, the obtained distance measures for a lot of combinations of physical parameters were correlated with the psychological distances of the perceptual dimensions extracted from perceptual ratings.

Having investigated a total of 43 combinations for two components and/or three components, 10 possible combinations showing the correlations between the perceptual dimensions and the acoustic and/or psychoacoustic magnitudes (see Table IV-6-1 to Table IV-6-9) deserved to be investigated, leaving us with such a combination of A-weighted pressure level, dB(A) Max for short, equivalent sharpness, Acum L_{eq} for short, and equivalent fluctuation strength, Vacil L_{eq} for short, which best account for the multidimensional properties of the perceptual solutions.

Table IV-6: Multiple correlations between psychological reference-axes and unidimensional components relating dB(A) Max, Acum L_{eq}, and Vacil L_{eq}

	dB(A) Max	**Acum L_{eq}**	**Vacil L_{eq}**
Sim_DimI	0.846**	–0.077	0.017
Sim_DimII	0.220	0.661*	0.134
Sim_DimIII	–0.218	–0.268	0.036
Pref_DimI	0.919**	0.043	0.220
Pref_DimII	–0.043	–0.351	–0.069
Pref_DimIII	–0.025	–0.371	–0.155

** correlation coefficient is significant on the level of 0,01 (two-tailed).
* correlation coefficient is significant on the level of 0,05 (two-tailed).

Table IV-6-1: Multiple correlations between psychological reference-axes and unidimensional components relating dB L_{eq}, Vacil Max, and Acum Max

	dB L_{eq}	**Vacil Max**	**Acum Max**
Sim_DimI	0.865**	-0.205	-0.076
Sim_DimII	-0.170	-0.545*	-0.397
Sim_DimIII	-0.277	-0.035	-0.234
Pref_DimI	0.703**	-0.369	-0.101
Pref_DimII	0.143	0.409	-0.384
Pref_DimIII	-0.307	-0.375	-0.457

Table IV-6-2: Multiple correlations between psychological reference-axes and unidimensional components relating Sone Max, Acum L_{eq}, and Vacil L_{eq}

	Sone Max	**Acum L_{eq}**	**Vacil L_{eq}**
Sim_DimI	0.788**	0.222	0.240
Sim_DimII	0.380	-0.656**	-0.291
Sim_DimIII	-0.232	0.219	-0.165
Pref_DimI	0.945**	0.075	-0.140
Pref_DimII	-0.098	0.341	-0.029
Pref_DimIII	-0.030	0.340	-0.224

Table IV-6-3: Multiple correlations between psychological reference-axes and unidimensional components relating dB(A) L_{eq}, Acum L_{eq}, and Asper L_{eq}

	dB(A) L_{eq}	**Acum L_{eq}**	**Asper L_{eq}**
Sim_DimI	0.819**	0.079	0.247
Sim_DimII	0.226	0.539*	-0.438
Sim_DimIII	-0.156	-0.370	0.001
Pref_DimI	0.910**	0.134	-0.257
Pref_DimII	-0.026	-0.358	0.154
Pref_DimIII	0.025	-0.327	-0.373

Table IV-6-4: Multiple correlations between psychological reference-axes and unidimensional components relating dB(A) L_{eq}, Sone L_{eq}, and Acum L_{eq}

	dB(A) L_{eq}	**Sone L_{eq}**	**Acum L_{eq}**
Sim_DimI	0.596*	0.534*	0.646**
Sim_DimII	0.622*	-0.342	-0.542*
Sim_DimIII	-0.265	-0.147	-0.054
Pref_DimI	0.894**	0.094	0.295
Pref_DimII	-0.243	0.131	0.336
Pref_DimIII	0.055	-0.486	-0.366

Table IV-6-5: Multiple correlations between psychological reference-axes and unidimensional components relating dB(A) Max, Acum L_{eq}, and Asper L_{eq}

	dB(A) Max	**Acum L_{eq}**	**Asper L_{eq}**
Sim_DimI	0.833**	0.092	-0.054
Sim_DimII	0.207	0.517*	-0.030
Sim_DimIII	-0.157	-0.368	-0.015
Pref_DimI	0.906**	0.130	-0.238
Pref_DimII	-0.016	-0.345	-0.084
Pref_DimIII	0.000	-0.355	0.133

Table IV-6-6: Multiple correlations between psychological reference-axes and unidimensional components relating dB(A) Max, Sone Max, and Acum L_{eq}

	dB(A) Max	**Sone Max**	**Acum L_{eq}**
Sim_DimI	0.768**	-0.146	0.325
Sim_DimII	0.446	0.060	-0.552*
Sim_DimIII	-0.260	0.190	-0.070
Pref_DimI	0.926**	0.160	0.038
Pref_DimII	-0.144	0.128	0.227
Pref_DimIII	-0.087	0.205	-0.529*

Table IV-6-7: Multiple correlations between psychological reference-axes and unidimensional components relating dB Max, Vacil Max, and Asper Max

	dB Max	**Vacil Max**	**Asper Max**
Sim_DimI	0.778**	0.155	0.183
Sim_DimII	-0.209	0.548*	0.076
Sim_DimIII	-0.017	0.038	-0.425
Pref_DimI	0.709**	0.238	0.349
Pref_DimII	0.279	-0.302	-0.060
Pref_DimIII	0.046	-0.177	-0.540*

Table IV-6-8: Multiple correlations between psychological reference-axes and unidimensional components relating Sone L_{eq}, Acum L_{eq}, and Asper L_{eq}

	Sone L_{eq}	**Acum L_{eq}**	**Asper L_{eq}**
Sim_DimI	0.744**	0.307	0.337
Sim_DimII	0.420	-0.669**	-0.107
Sim_DimIII	-0.200	0.223	-0.268
Pref_DimI	0.939**	0.011	-0.035
Pref_DimII	-0.131	0.398	-0.036
Pref_DimIII	0.071	-0.006	-0.584*

Table IV-6-9: Multiple correlations between psychological reference-axes and unidimensional components relating Sone Max, Acum L_{eq}, and Asper L_{eq}

	Sone Max	**Acum L_{eq}**	**Asper L_{eq}**
Sim_DimI	0.800**	-0.104	0.183
Sim_DimII	0.354	0.517*	-0.185
Sim_DimIII	-0.187	-0.309	-0.186
Pref_DimI	0.941**	0.058	-0.171
Pref_DimII	-0.063	-0.336	-0.078
Pref_DimIII	-0.008	-0.281	-0.273

A short note on the dB(A) metric: Since 1967, “lower is better” has been the motto in sound quality design that relied almost entirely on sound-power response data. The A-weighted sound level has widely been used as a predictor of acceptability or preference of automotive sounds. For dB(A), the A-weighted level versus time is calculated using three time constants “impulse” (L_{AI} = 35 ms), “fast” (L_{AF} = 125 ms) or “slow” (L_{AS} = 1000 ms). In practice, L_{AF} is commonplace as the perceived loudness remains constant for sounds with durations larger than 100 ms (Fastl, 1997). It is known that the maximum time resolution of the human hearing system is about 200 ms (ANSI S1.11; Zwicker & Fastl, 1999). A drawback of A-weighting that it underestimates the psychoacoustic measure of loudness at low frequencies below about 200 Hz, hence dB(B), dB(C) have been developed. For the case in which the inside noise of motor vehicles involves low frequency components below about 100 Hz, for instance, the B-weighted sound level, dB(B), is recommended. Thus from the difference, dB(B) – dB(A), the effect of low frequency components upon overall noise level can be estimated (DIN 45 639). For the purpose

of aurally adequate assessments of air traffic noise, for instance, Kryter (1967) has introduced the concept of dB(D). A general review on spectral weightings of the SPL measurement can be found in the paper by Schick (1990b).

Table IV-5: Psychological distances interrelating the 15 sounds on the basis of three unidimensional components dB(A) Max, Acum L_{eq}, and Vacil L_{eq}

Dissimilarity measure: Euclidean metrics

$d_i \backslash d_j$	GG	AVG	VMG	ASG	DG	TGG	KG	BQG	W1G	W2G	R1G	R2G	KAG	IBG	GHG
GG	0														
AVG	26.55	0													
VMG	14.75	36.28	0												
ASG	30.06	6.018	38.93	0											
DG	21.91	10.33	28.64	10.69	0										
TGG	31.64	5.204	40.83	4.598	13.36	0									
KG	34.14	10.23	42.23	4.528	13.60	6.81	0								
BQG	22.05	32.91	35.87	37.15	34.89	37.4	41.49	0							
W1G	33.14	6.851	42.91	5.718	15.51	2.711	7.253	37.44	0						
W2G	23.49	10.65	36.20	13.29	15.76	14.13	17.32	24.46	13.83	0					
R1G	36.96	11.89	44.51	7.775	16.10	7.341	4.683	44.41	7.979	20.55	0				
R2G	21.57	27.51	36.19	31.03	29.93	31.68	35.16	7.88	31.45	17.93	38.41	0			
KAG	34.47	8.922	42.39	7.94	15.05	4.886	7.954	41.54	6.535	18.86	5.453	36.16	0		
IBG	33.19	7.896	41.58	9.175	15.36	5.155	10.46	39.53	6.827	17.85	8.618	34.51	3.25	0	
GHG	23.4	9.659	29.81	10.04	2.161	12.26	12.74	36.15	14.62	16.56	14.71	31.27	13.4	13.74	0

Table IV-7: Psychological distances interrelating the 15 sounds according to perceptual ratings and physical measures

stimulus	**component distances X_{dBA}**	**component distances X_{Acum}**	**distances of first similarity dimension Sim_DimI**	**distances of second similarity dimension Sim_DimII**	**distances of first preference dimension Pref_DimII**	**distances of second preference dimension Pref_DimII**
gg	-1,22	-,41	-,80	-,95	-1,74	,31
avg	,28	,08	,73	,33	,41	-,25
vmg	-1,40	-1,26	-,69	-1,15	-1,37	-,82
asg	,51	,02	,64	-,33	-,15	-,25
dg	,07	-,44	-,52	-,85	,29	,17
tgg	,58	,12	,84	-,29	1,12	-,23
kg	,75	-,02	,48	-,01	,13	,01
bqg	-1,51	,86	-1,47	,80	-1,40	,32
w1g	,63	,27	,57	,35	,65	-,49
w2g	-,14	,47	,29	,09	-,46	,59
r1g	,95	-,07	,95	,22	,87	,21
r2g	-1,13	,86	-1,23	,60	-,80	-,28
kag	,79	-,04	,97	,34	,86	,17
ibg	,67	,03	-,02	-,01	1,05	-,20
ghg	,17	-,46	-,75	,85	,54	,74

Table IV-6 gives the multiple correlations between the coordinates on the six perceptual dimensions and those on the three component dimensions. One notes in this table that the correlation between the first dimension of similarity judgments (Sim_DimI) and A-weighted maximum pressure level is greater

than 80%, between the second dimension of perceived similarity (Sim_DimII) and sharpness in acum greather than 60%, whilst the correlation between the first dimension of preference ratings (Pref_DimI) and A-weighted maximum pressure level is greather than 90%. Since the three-dimensionality for the representation of perceptual structures is unnecessary, the number of component dimensions here is restricted to two, hence sound stimuli, varying in dB(A) Max and Acum L_{eq}, were investigated in detail. The three physical parameters were analyzed by Kruskal's method (see Table IV-5).

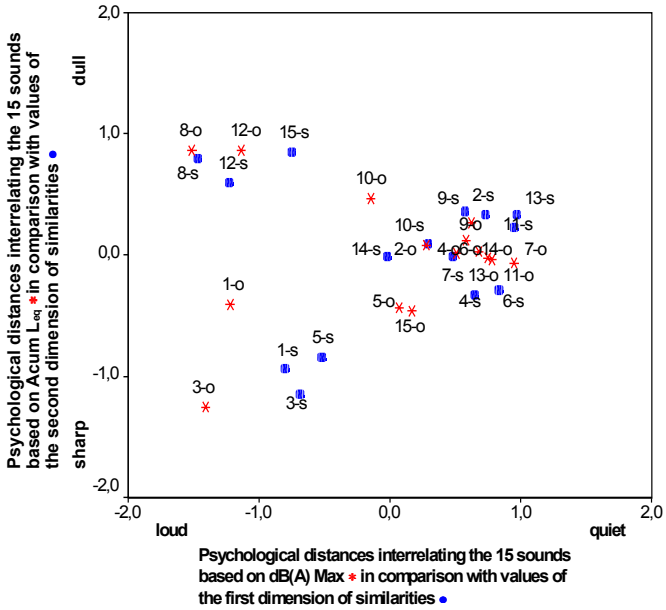

Fig. IV-1: Arrangement of stimuli regarding composite distances (•) based on dimensions of perceived similarity relative to its component distances (*) on the basis of dB(A) Max and Acum L_{eq}

Both subjective scores and physical values are written as metric distances (see Table IV-7). We are now left with a common stimulus space of lower dimensionality in which two orthogonal axes refer to dB(A) Max and Acum L_{eq} respectively. Suppose now that the perceptual data are plotted against two physical parameters, measured for every sound in the set, we would then construct a multidimensional space, while its coordinate axes representing the component dimensions in terms of A-weighted maximum sound level and sharpness. In Fig. IV-1, Fig. IV-2 the set of sound samples is represented as a set of stimulus points in the composite multidimensional similarity as well as preference space aligned with physical variables under study. Two orthogonal component dimensions span up the common stimulus space onto which even

two perceptual dimensions pertaining to preference ratings and perceived similarity are projected. It has the advantage that all the data can be presented in a single geometrical framework. Letters designated with the symbol * indicate the positions of sound samples determined objectively, letters tagged with the symbol • are assigned arbitrarily to the 15 stimuli for subjective positions, correspondingly. With high correlations, as is shown in Table IV-6, we have plotted the unidimensional component distances, one for sound level and the other for sharpness, simultaneously the psychological distances underlying the perceptual dimension pertaining to perceived similarity as abscissa against the distances underlying the second similarity dimension as ordinate. In ideal case, the positions of subjective points would be perceived at exactly the same locations of physical points. In view of the fact that there is some continuous overlapping of distributions of perceptual points and objective ones, it is apparent from the composite configuration that sound level and sharpness are considered to be meaningful substantive features shared by stimuli. In case of perceived similarity, the subjective R2G (12-S for short) as well the subjective BQG (8-S for short) are positioned in the vicinity of the objective R2G (12-O for short) as well the objective BQG (8-O for short), hence R2G and BQG are perceived to be loud and dull both subjectively and objectively.

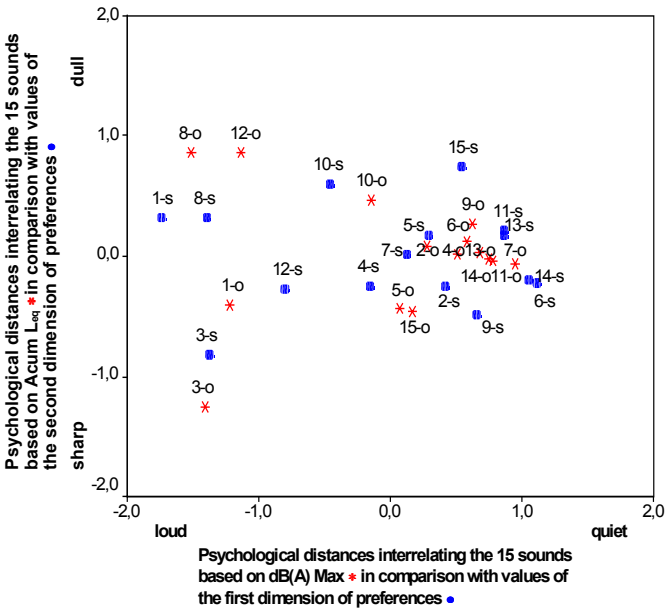

Fig. IV-2: Arrangement of stimuli regarding composite distances (•) based on dimensions of preference ratings relative to its component distances (*) on the basis of dB(A) Max and Acum L_{eq}

Contrary to BQG, the positions of 12-O and 12-S in the preference structure show a gross departure to the degree that both 12-S and 12-O are perceived to be loud but 12-S approaches the area of being sharp, which counters the fact that 12-O is objectively declared as dull. The objective GHG ('15-O') is measured as sharp, and the position of 15-O lies about midway between 15-S in the similarity structure and 15-S in the preference structure. 15-S is perceived in both cases to be dull, varying largely in sound level. Considering GG ('1-O') and DG ('5-O') in the composite scaling space of similarities and physical measures, due to its statistical significance, 1-S and 5-S are given much more emphasis on their sharpness. With reference to the sound level, 1-S in the preference structure is considerably greater than 1-O. It is apparent from the composite space of similarities and physical quantities that the sharpness factor is considered to be a psychologically meaningful attribute in the multidimensional property for BQG ('8-S'), R2G ('12-S'), GG ('1-S'), VMG ('3-S'), and DG ('5-S'). The subjective GHG ('15-S') and the objective one ('15-O') differ in sound level as well as in sharpness. Despite GHG's position ('15-S') being near BQG ('8-S') and R2G ('12-S'), an additional factor of fluctuation strength might be suspected, however for there being statistically nonsignificant, we have not considered in interpretations of the data.

4.3 Physical Basis of Perceptual Dimensions

It would be expected that the common stimulus space of the measured sounds would contribute to the functional interpretations of the perceptual data. Hence it is of great interest to closely investigate the extent to which the psychological distances with respect to the locations of the judged stimuli will deviate from the locations of the ones on the basis of two physical parameters under consideration. Plotting the psychological distances among the sound samples relating to the coordinates of perceptual ratings and its physical counterparts against the differences in sound level and sharpness, we observe that with respect to sound level in the similarity case (see Fig. IV-3) the empirical points IBG ('14-S') and GHG ('15-S') lie far off the reference line, marked by a solid line, suggesting that those stimuli spreading about the empirical line, marked by a dashed line, are appreciably larger than expected. The regression coefficients give us –0.853 for a hypothetical function ($p < .01$) as well as –0.736 for an empirical one ($p < .05$), respectively. With the values for IBG ('14-S'), GHG ('15-S'), and DG ('5-S'), we may infer that 14-S lies 0.69 units away from IBG-O on a unique dimension of its own, namely dB(A) Max, likewise 15-S deviates from 15-O by 0.92 units, and 5-S deviates from 5-O by 0.59 units.

Table IV-8: dB(A) Max and Acum L_{eq} interrelating the 15 sounds and estimated subjective parts for perceived similarity and preference ratings

stimuli	dB(A) Max	Acum L_{eq}	estimated dBA Max for Sim_DimI by 0.35 units (X ± 2)	estimated dBA Max for Pref_DimI by 0.29 units (X ± 1.66)	estimated Acum L_{eq} for Sim_DimII by 0.33 units (X ± 2.14)	estimated Acum L_{eq} for Pref_DimII by 0.48 units (X ± 2.4)
gg	81,00	3,16	78,60	83,98	6,66	0,44
avg	67,30	1,56	64,73	66,56	-0,06	3,21
vmg	83,20	3,82	79,14	83,03	3,11	1,62
asg	72,00	1,30	71,16	75,78	3,57	2,65
dg	75,40	1,96	78,77	74,14	4,62	5,01
tgg	65,70	1,25	64,21	62,61	3,91	3
kg	72,50	1,07	74,04	76,05	1,01	0,92
bqg	78,10	2,75	77,87	77,47	3,14	5,45
w1g	66,20	1,11	66,54	66,08	0,59	-2,69
w2g	75,80	1,64	73,34	77,63	4,1	2,24
r1g	66,70	,98	66,70	67,16	-0,9	-0,42
r2g	81,80	2,32	82,37	79,91	4,01	8,02
kag	61,40	1,20	60,37	61,00	-1,26	0,15
ibg	58,80	1,31	62,74	56,62	1,57	2,46
ghg	72,80	1,91	78,06	70,68	-6,59	-4,09

The psychological distances of some empirical points, such as R1G ('11-S'), R2G ('12-S'), BQG ('8-S') and W1G ('9-S'), are almost identical to its physical counterparts. In addition, the distance of GG ('1-S') from GG ('1-O') or of AVG ('2-S') from AVG ('2-O') is shown to be twice the distance of TGG ('6-S') from TGG ('6-O') or of KG ('7-S') from KG ('7-O'). The psychological differences between the empirical and the physical stimulus points render the value of 0.35 on average, which approximately amounts to 2 dB(A) Max. This gain of 2 dB(A) Max can be adopted under the assumption that the theoretical function is linear on the basis of physical stimulus points. By the gain of 2 dB(A) Max all sound samples have been estimated in order to make a possible functional relationship between the MDS solution pertaining to the similarity-difference judgments and its acoustic counterpart. The subjective values with reference to dB(A) Max compared to the physical parts can be estimated from the distance differences and have been tabulated in Table IV-8 under the heading "X ± 2". Fig. IV-4 shows the predicted similarity judgments for sound stimuli in terms of A-weighted maximum pressure level. The deviations of the values from the prediction show some consistency. This result can be taken as an indication that the predicted

values, as measured in dB(A) Max, determine the empirical values by raising 0.35 units in distance measures.

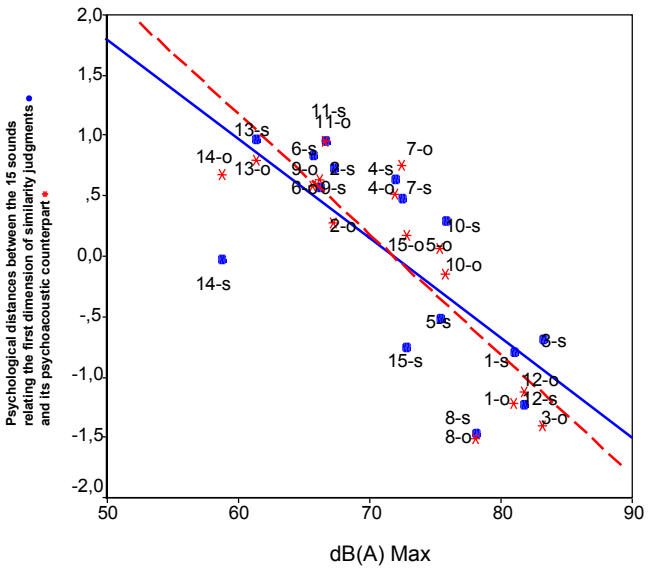

Fig. IV-3: Distances between the objects relating the first dimension of perceived similarity (•) and its physical component (*) as a function of dB(A) Max

Plotting the corrected subjective values against the physical measures, we get a simple linear relationship, a straight line with slope. The range of variation on which the values are based is limited between 55 dB(A) Max and 85 dB(A) Max. If this approximation on the abscissa is appropriate, in other words, if similarity judgments are to have a particular psychological framework which would be logically adequate for a description of the acoustic parameter, it is possible to write a modified equation of the form

$$X = a + b\,[dB(A)\ Max] + G \qquad (9.1)$$

in which a and b are constants, both of which are estimated by Least Squares Regression [26] to analyze the data with respect to perceptual judgments and an acoustic parameter, and G is a gain factor adjusting the regression line. Taking 0.35 units difference into account, the relation between the predicted

[26] In simplest form, Least Squares Regression can be written as follows:

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \min. \qquad (9.0)$$

where the regression line minimizes the squared sum of standard errors. Least Squares Regression is thought of as a multivariate technique and makes it possible to understand the relationship between the subjective values and the objective measures (e.g. Bortz, 1993, 170).

values of X and the actual judgments can be approximated by a straight line:

$$X_{sim} = 7.16 + 0.9\ [dB(A)\ Max]\ \pm\ 2 \qquad (9.2)$$

The error was found to be negligible (< 1%) thus not considered in interpretations of the data. The very high multiple R associated with the regression is 94% ($F(1,13) = 92.38$, $p < .01$). The corrected multiple R (corrected $R^2 = 0.87$), sometimes called the determination coefficient as a measure of correlation performance between subjective evaluations and physical magnitudes, quantifies the Least Sqaures fit of a line to the data. Due to a high correlation, 87% of the variance of the subjective values is accounted for by a loud-soft scale. Much of the variability about the curve is caused by some stimuli, which show more deviations than we should anticipate from the general upward trend. Sound sources aligned with GHG, DG, IBG, R2G, and KG which are positioned above the line can be overestimated rather than underestimated.

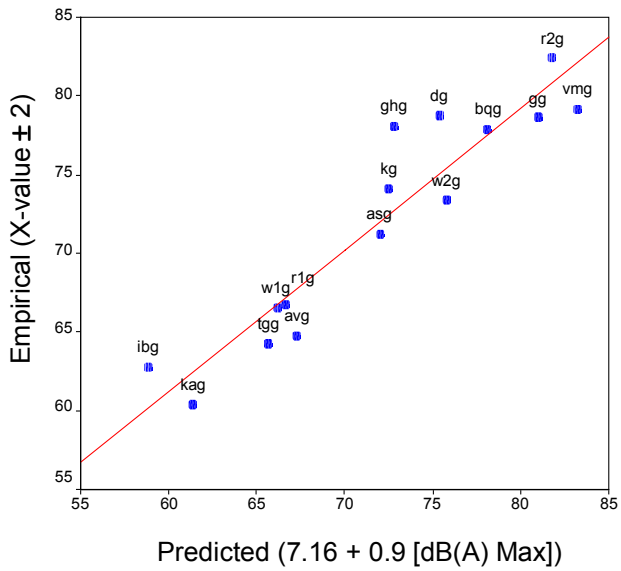

Fig. IV-4: Prediction of psychological differences between sounds in similarity-difference judgments from optimal weighting of physical variables

In similar fashion we proceed to the case where the distances between the empirical stimulus objects on the basis of the like-dislike ratings and corresponding physical ones are plotted against the difference in terms of sound level. The results show the same effect. From Fig. IV-5 we find that the empirical regression line with the coefficient of –0.890 ($p < .01$) represented by a dashed line is nearly overlapping with the reference line represented by a

solid line, the coefficient of which is calculated as being –0.853 ($p < .01$). In 11 of the 15 cases excluding GG (‘1-S’), ASG (‘4-S’), W2G (‘10-S’), and KG (‘7-S’), the most empirical stimuli are somewhat smaller than the hypothetical ones should be, with respect to sound level. The discrepancies of those empirical stimulus points like VMG (‘3-S’), W1G (‘9-S’), R1G (‘11-S’), KAG (‘13-S’), IBG (‘14-S’), and GHG (‘15-S’) from their physical counterparts are almost identical. Regarding the value for 10-S, we may infer that 10-S lies 0.32 units away from 10-O, likewise 4-S deviates from 4-O by 0.66 units, and it is twice as much as the metric distance between R2G (‘12-S’) and R2G (‘12-O’).

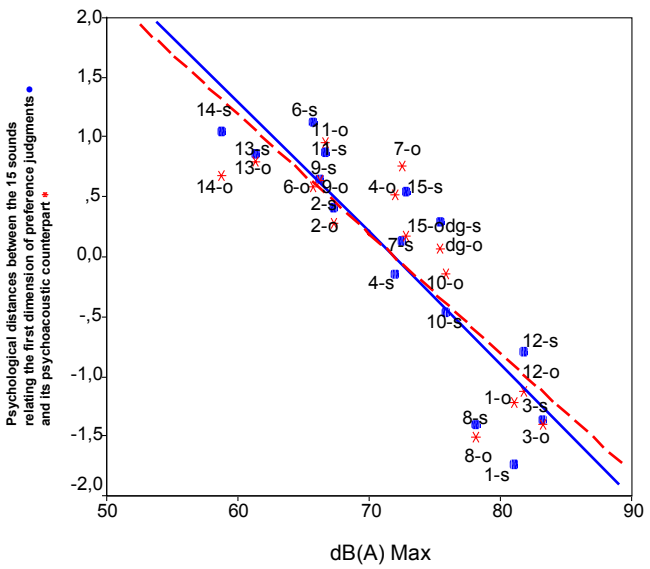

Fig. IV-5: Distances between the objects relating the first dimension of preference ratings (•) and its physical component (*) as a function of dB(A) Max

The mean distance disparity between the empirical and the physical stimulus points is calculated as being 0.29, corresponding to approximately 1.66 dB(A) Max. The subjective values in dB(A) Max compared to the physical parts have been estimated and are listed in Table IV-8 under the heading “X ± 1.66”. Assuming the linearity of the theoretical function the gain of 1.66 dB(A) Max can be adopted for estimating the extent to which the sound samples in like-dislike ratings go along with the unidimensional component. The curve in Fig. IV-6 shows the predicted preference ratings for sound stimuli in terms of A-weighted maximum sound level. The range of variation on which the data is based is limited between 55 dB(A) Max and 85 dB(A) Max. Likewise, when we plot the estimated values as a function of differences in terms of dB(A) Max, as the case may be, the resulting function appears to

be reasonably linear. Taking 0.29 units differences into account, the prediction model of subjective preference in terms of linear regression is defined by:

$$X_{pref} = -5.51 + 1.08\ [dB(A)\ Max] \pm 1.66 \qquad (9.3)$$

The error was shown to be less than 1% thus it was excluded from the model specification. The very high multiple R associated with the regression is 97% ($F(1,13) = 195.45$, $p < .01$). The correction of the multiple R (corrected $R^2 = 0.93$). Due to a strong correlation, 93% of the variance of the subjective values is accounted for by loud-soft scale. The unexplained 7% may be subjected to some objects which are slightly deviant or attributable to a scaling error. Recall that the instrumentally measured values of sounds have been converted to the values in percentage!

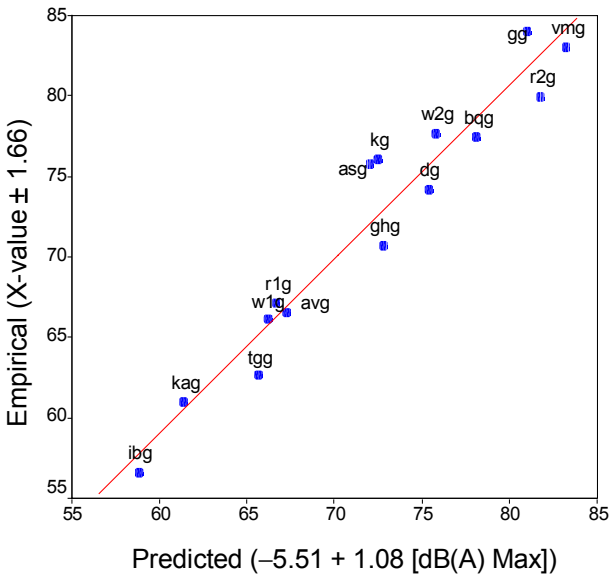

Fig. IV-6: Prediction of psychological differences between sounds in like-dislike ratings from optimal weighting of physical variables

As a consequence it predicts that those sound sources like ASG, KG, and GG which are positioned above the line can be overestimated rather than underestimated. Anyhow, the curve displays a much better Least Sqaures fit of empirical and predicted values than has been the case of similarity. The scatter of stimulus points does carry a suggestion of the general upward trend, i.e., acoustic parameter A-weighted maximum sound level is a clearly correlated factor for preference ratings.

Much the same analysis holds for the case with regard to the equivalent sharpness, however the relationship among the pairs of stimuli varying along the sharpness continuum leaves us unimpressive results to be seriously considered. We find that the deviations from the prediction show no particular consistency (see Fig. IV-7): With reference to equivalent sharpness the curves representing the Least Sqaures fit between the predicted values based upon the second dimension pertaining to perceived similarity and the measured Acum L_{eq} of the sounds, with the correlation of 39% ($F(1,13) = 2.29$, $0.15 > p = .1$), show no statistical significance. It is obvious that the variation in dB(A) Max is psychologically much more important than the variation in Acum L_{eq}. Much the same holds for the preference case.

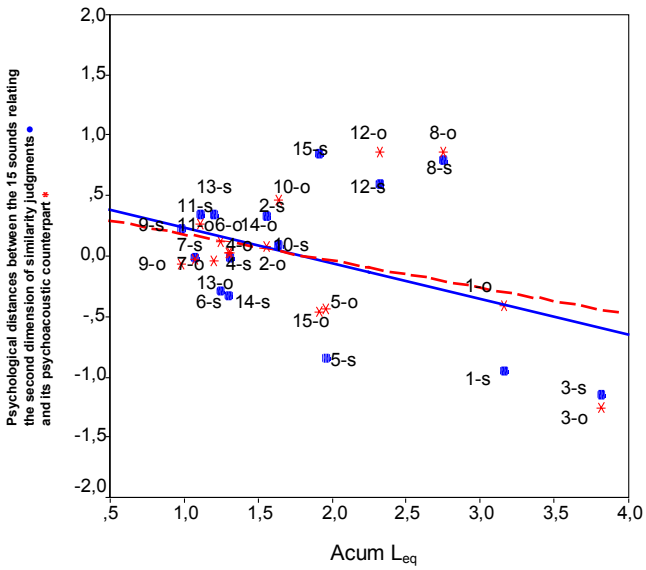

Fig. IV-7: Distances between the objects relating the second dimension of perceived similarity (•) and its psychoacoustic counterpart (*) as a function of Acum L_{eq}

The parameter Acum L_{eq} appears to have only a slight effect on preference ratings, showing the correlation of 25% ($F(1,13) = 0.86$, $0.37 > p = .1$) with regard to the second preference dimension. The mean disparity between the multidimensional distances based upon the second similarity dimension and corresponding component distances gives us 0.33, which corresponds to approximately 2.14 Acum L_{eq}. The subjective distance values in reference to the equivalent sharpness have been corrected by a factor of 2.14 and tabulated in Table IV-8 under the heading "X ± 2.14". Analysis of the subjective distances based on the second similarity dimension, corrected by a factor of 0.33, and the component distances based on Acum L_{eq} of the sounds render the relatively high multiple R of 68% ($F(1,13) = 11.34$, $p < .01$), which has

statistical significance. Even though having corrected the distance values of the second preference dimension by a factor of 0.48, it revealed the correlation of 24% (F(1,11) = 0.65, 0.44 > p = .1), showing statistical nonsignificance.

4.4 Additive Similarities Model

Sound evaluation is viewed as an unitary percept resulting from an integration of various prevailing physical factors. The representation of the stimuli as a set of elements, aspects, or features can be thought of as a sort of associationistic approach. Tackled with the problem of how disparate pieces of information of stimulus objects in interest are combined or integrated to release the overall response, Anderson (1973a,b,c) has offered algebraic models by which complex cognitive activities can be explained quantitatively, giving a following note that *"The organism is considered as an integrator of stimulus information, not as a conditioning machine." (Anderson, 1974, 266)*. His information integration theory for psychological measurement of social objects (e.g. personality profiles, leaderships etc.) involves models of adding, averaging, subtraction, multiplying, or ratios. With regard to the ranking order or choice of data, two models of averaging and adding are of special importance. Since the basic ideas are simple, only a few sketches will be made in the following section. Consider two stimuli, i, j, with the scale values, S_i, S_j, and weight W_S on dimension S, etc. Perceived values can be viewed as the locations of stimuli along the dimensions. Each stimulus object is allowed its own weight parameter. Weight may be viewed as the amount of information in the stimulus. The averaging model, which underlies the city-block model, specifies the overall response as a weighted mean of the specific distances among two stimulus points being compared, by definition:

$$R_{ij} = C + \frac{W_S d_S(S_i, S_j) + W_L d_L(L_i, L_j) + \cdots}{W_S + W_L + \cdots} + \varepsilon_{ij} \tag{9.4}$$

where C is an additive constant allowing an arbitrary zero in the response scale, and will be assumed to be zero. The error term ε is an additive random variable with a mean of zero, thus generally ignored. The denominator forces the relative weights to sum to one. One restriction calls for equal weighting of the stimulus coordinates along a given dimension of the solution, which is necessary for the parallelism prediction. With differential weighting, the averaging model becomes nonlinear. Consider two ordinal response data, R^1, R^2, for simplicity, along two dimensions, S, L, the prediction of parallelism under the averaing model should be possible:

$$R_{ij}^{1} - R_{ij}^{2} =$$
$$[(W_S^1 d_S(S_i^1,S_j^1) + W_L^1 d_L(L_i^1,L_j^1)] - [(W_S^2 d_S(S_i^2,S_j^2) + W_L^2 d_L(L_i^2,L_j^2)]. \quad (9.5)$$

Setting the sum of weights equal to 1, the averaging model reduces to the simple adding model. Little is known which algebraic integration model, if any, might underlie the similarity judgments on vehicle interior noise. A priori, there is no great reason to believe that pairwise similarity judgments will be based on an adding model. An adding model, however, could be an attractive possibility if the assumption of stimulus independence is not required. Along with Beals et al. (1968), Tversky and Krantz (1970, 1975), or Gati and Tversky (1982), two properties of perceptual dimensions that provide information about the manner on how subjects integrate component dimensions in a perceptual task: First, interdimensional additivity asserts that the contributions of different component dimensions to overall similarity are combined in such a manner that overall similarity is monotonically related to the sum of the perceived values on a set of component dimensions to be judged separately. Secondly, intradimensional subtractivity states that the contribution of any component dimension to overall similarity depends on the absolute subjective differences among the two percepts along that component dimension.

Consider combinations of two stimuli i, j, each defined along several component dimensions S, L,. . . , and let d_S be the similarity function on a component dimension S, etc., so that the stimulus i can be represented as (S_i, L_i, . . .), where S_i is its perceived value on the component dimension S, etc. No metric is assumed for these perceived values which may be discrete and nominal. The overall similarity between the stimuli i and j can be written as an interdimensional integration rule

$$R_{ij} = W_S d_S(S_i,S_j) + W_L d_L(L_i,L_j) + \cdots \quad (9.6)$$

where the W's are weighting coefficients reflecting the relevance of component dimensions in the similarity judgments. If differential weight estimates are to be anticipated, then it may be advisable to get a separate index of the weight. It is assumed that interdimensional similarity in terms of the cognitive operation of integration (e.g. Schroeder et al., 1967) follows an adding model, $d_S(S_i,S_j) = (s_i + s_j)$, whereas intradimensional similarity in terms of the cognitive operation of discrimination (e.g. Schroeder et al., 1967) follows an difference model, $d_S(S_i,S_j) = |s_i - s_j|$. Here s_i and s_j represent the similarity ratings on the component dimension S. Factor analyzing the scalar products from similarity of pitch (e.g. Eisler & Ekman, 1959), or from similarity of emotions (e.g. Ekman et al., 1964), one has a nonlinear similarity function, $R_{ij} = s_i/(s_i + s_j)$, $s_i \leq s_j$, where i, j are two percepts, and their subjective magnitudes, s_i, s_j, and their similarities R_{ij} on a scale ranging from 0 ('dis-

similar‘) to 1 (‘identical‘). As such, perceptual similarities of pitch as a function of centre frequency can be equivalently expressed as scalar products, $b_{ij} = \Sigma_k x_{ik} \cdot x_{jk}$, equivalent to the squared Euclidean distances, $d^2_{ij} = \Sigma_k (x_{ik} - x_{jk})^2$, between the origin (the lower pitch) and the endpoints of two vectors (the mean of two pitches), computed by using the cosine law, $d^2_{ij} = h^2_i + h^2_j - 2h_i h_j \cdot cos(\alpha)$. Thereby h_i, h_j correspond to the vector lengths **i**, **j** from the origin, and α is the angle subtended by **i** and **j**. The overall similarity is supposed to be a function of common stimulus dimensions with the lengths of two vectors, thus different vector lengths represent their intensities, whereas different directions of the vectors represent their qualitative differences (for a more details see Borg & Groenen, 1997, Chapter 17).

Attneave (1950) first carried out a metric test of an additive difference model mediated by a rule of combination (e.g. r = 2/3) on how differences along two or more underlying stimulus dimensions combine to yield the overall similarity among the two stimulus objects. The additive difference model assumes that similarity judgments between multidimensional objects are regarded as composed of two independent processes, i.e. an interdimensional additive process and an intradimensional difference process, by definition (Beals et al., 1968, 134; Tversky & Krantz, 1970, 574):

$$d(x,y) = F\left[\sum_{i=1}^{n} \phi_i\left(\left|X_i - Y_i\right|\right)\right] \tag{9.7}$$

where both F and the ϕ_i are strictly increasing functions of one variable and $X_i = f_i(x_i)$ is the perceived value of stimulus object x on the i-th dimension. The same holds for Y_i. After having studied two hypotheses relating to the similarity function, one for Euclidean hypothesis (Richardson, 1938) and the other for additive hypothesis (Householder & Landahl, 1945), Attneave (1950) has picked up the latter and applied it to his own data. According to Richardson, two color patches in a 2D configuration differ by a units on one sensory attribute, called “brightness”, and b units on another sensory attribute, called “saturation”, which are represented by a set of mutually orthogonal axes in an Euclidean space, i.e. subjectively independent. Householder and Landahl made a supposition that the subjective differences on the stimulus dimensions combine additively, whereby composite similarity judgments would be predicted by a simple or weighted linear relationship.

Such an adding model allows us to account for the perceptual dimensions revealed in preference as well as similarity studies by estimating the perceptual weights assigned to each perceptual dimension and the set of acoustic and psychoacoustic magnitudes by each sound. Hence it can be postulated that perceived similarity is to be summed up across a set of perceptual dimensions extracted from the multidimensional representations. One hypothesis we wish to test is the overall similarity is an additive function

of two physical components. For the following discussion it is assumed that once there is some consistent suggestion of linearity in the relationship between the composite MDS solutions and the sum of unidimensional component distances, thus the application of the adding model to the similarity judgments should be possible. Taking this into account, a relatively simple formula can be given for the additive similarities model of sounds:

$$X_{sum} = C + W_1 [X_{dBA}] + W_2 [X_{Acum}] + G \qquad (9.8)$$

where X_{sum} is the predicted total similarity, X_{dBA} and X_{Acum} are the measured distances on the physical scales, dB(A) Max and Acum L_{eq}, respectively. The so-called weighting coefficients W_1, W_2 are the projections of X_{sum} on the orthogonal component dimensions. C, W_1, and W_2 are to be estimated by Least Squares Regression (e.g. see Diederich et al., 1957) to analyze the distance values of X_{sum} with respect to similarity judgments of sound samples and two physical parameters. The regression weights suggest a particular calibration of the two component dimensions. The additive constant G is not a free parameter, but rather a fixed value that must be estimated from the data. It allows the weighted values to have a natural zero corresponding to a line of zero slope. The comparison of the relevance of two qualitatively different sensory attributes is not allowed since there is no way to distinguish between a sensory attribute that is effective by virtue of a large range in distance values, and a sensory attribute that is effective by virtue of a high weight. Differential weighting of the one attribute will depend on the other attribute with which it is combined.

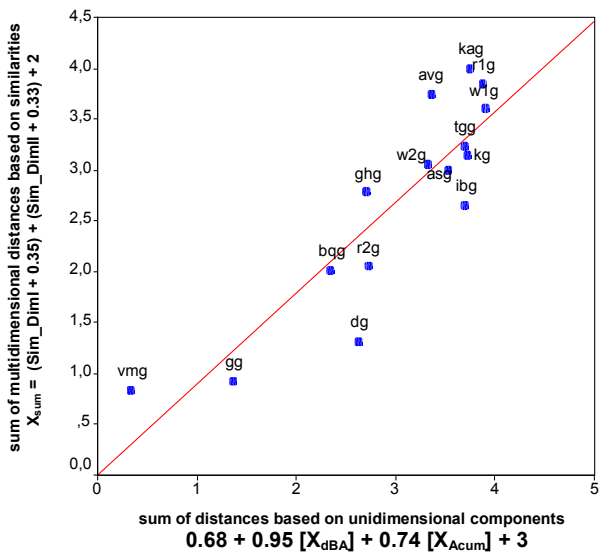

Fig. IV-8: Total of each weighted multidimensional scales as a function of sum of unidimensional projections

In accord with the adding hypothesis, the addition of Acum L_{eq} thus increases perceived similarity. Whether it is meaningful to add X_{dBA} and X_{Acum} depends not on the physical properties of sounds but on the uniqueness properties of the numerical assignments associated with X_{dBA} and X_{Acum}. This comparison is made in Fig. IV-8. The integration model rests on the ground that if perceived similarity is to be summed up of dB(A) Max and Acum L_{eq}, then the similarity-component relation, i.e., relation between the sum of psychological distances and the sum of component distances should be linear with zero intercept. The ordinate for each stimulus point is the sum of two unidimensional component distances, one for dB(A) Max and the other for Acum L_{eq}. The psychophysical map underlying a linear model is defined by two reference points and a slope parameter. The two reference points are a zero point. To compare the MDS solutions with the unidimensional components, the zero-points of the two scales would be corrected in such a manner that the lowest of distance values is taken as a zero-point, by adding 2 units for perceptual dimensions and 3 units for component dimensions. The two MDS solutions have been corrected by a factor of 0.35 and 0.33 respectively, both of which is found to show a degree of internal consistency. The prediction model in terms of linear regression is defined by

$$X_{sum} = 0.95\ [X_{dBA}] + 0.74\ [X_{Acum}] + 0.36 \qquad (9.9)$$

The psychological differences on two physical components are combined to an overall similarity by means of an additive constant having a fixed value 0.36 that could be estimated from the data. Each perceptual dimension has specific weight, i.e. two stimuli in a 2D space differ by weighting coefficients of 0.95 on X_{dBA} and of 0.74 on X_{Acum}, respectively. The interrelatedness between X_{dBA} and X_{Acum} would correspond to the W-parameters of Eq. 9.8, which act as a multiplier of the perceived values of the physical components. The correlation here is 89% ($F(2,12) = 21.89$, $p < .01$). The very high multiple R associated with the regression indicates that the overall similarity predicts 89% of the variance in the data. The slope of the function is the relevance of component dimensions to sensory attributes pertaining to perceived similarity. The correction of the multiple R of 75% suggests that trends exist. So perceived similarity as a function of dB(A) Max and Acum L_{eq} (expressed in terms of estimated distance values) can be approximated by an adding model. This prediction model developed by Least Sqaures Regression confirms a better graphical fit than an alternative model like the multiplying model, but no test of goodness-of-fit is possible.[27]

[27] To get a proper test of fit, a logarithmic transformation of the distance values was attempted with the present data, on the assumption that taking logs will convert the two-factor multiplying model to an adding model if C = 0. The outcome was unsatisfactory. Such a log-method disallows negative numbers. On the other hand, correlation coefficients for testing fit are sometimes thought to be inappropriate. This topic is discussed further in Birnbaum (1973).

With regard to the effect of dB(A) Max and Acum L_{eq}, Fig. IV-9 further indicates that pairwise similarity judgments for those stimuli aligned with GG, VMG, BQG, and R2G seem to be facilitated by a distinctive feature Acum L_{eq}. Why the stimulus object tagged with capital letter DG lies far apart from the additivity hypothesis (see Fig. IV-8) we do not know. This deviation seems to be attributable to the effect of an additional component Vacil L_{eq} but this hypothesis does not find support in our data.

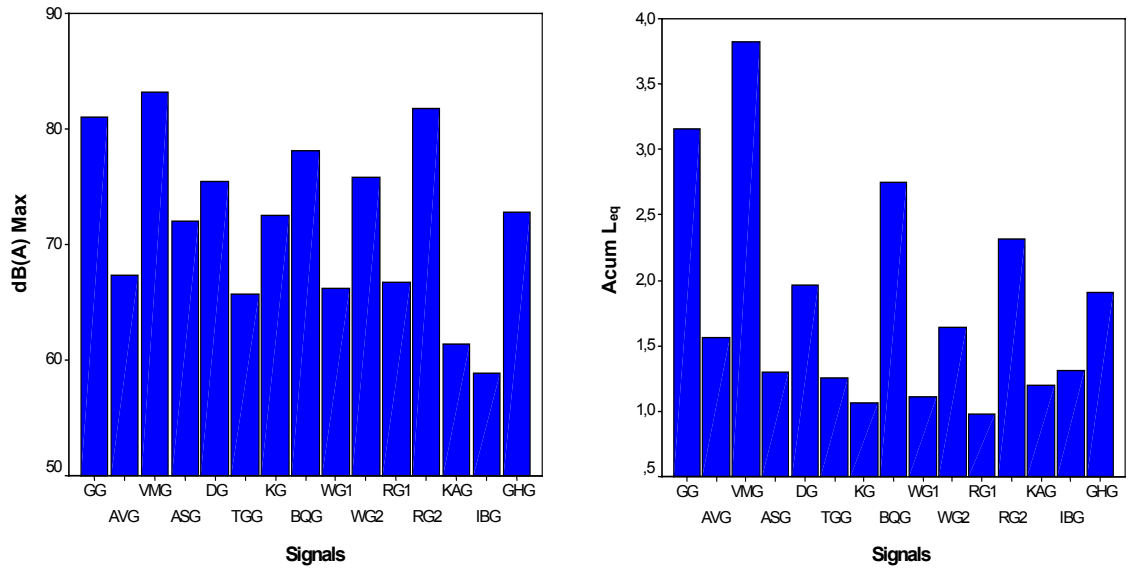

Fig. IV-9: dB(A) Max and Acum L_{eq} of sounds in percentage

The psychological and physical scales appear to be monotonously related. Using the data available, we could demonstrate that perceived similarity is approximated as a weighted sum of dB(A) Max and Acum L_{eq}. It is concluded that the adding model can provide a quantitative account of "level-sharpness" effect, in other words, when judging the perceived similarity among the two sounds, subjects make direct use of two disparate pieces of stimulus information. The same stimulus will have different relevance and different valence for different judgment tasks. By virtue of which the physical parameters are defined, the integration function will be taken for granted. The conclusion is clear: The adding model gives a good account of the similarity data, both conceptually and in quantitative form. There is no definite answer available from any other data confirming that perceived similarity of vehicle interior noise be found to correlate highly with sound level and sharpness. Therefore, the approximation based on Eq. 9.9 would be limited in its applicability.

4.5 Discussion

Nonmetric multidimensional scaling techniques were used to capture the common perceptual dimensions shared by a set of vehicle interior noises, and perceptual features specific to each sound. The multidimensional analysis of the data allowed a projection of the perceptual structures underlying the judgments of preference and perceived similarity onto a common stimulus space. In a multiple correlation analysis, a number of acoustic and psychoacoustic magnitudes were shown to have good correlations with the perceptual dimensions extracted.

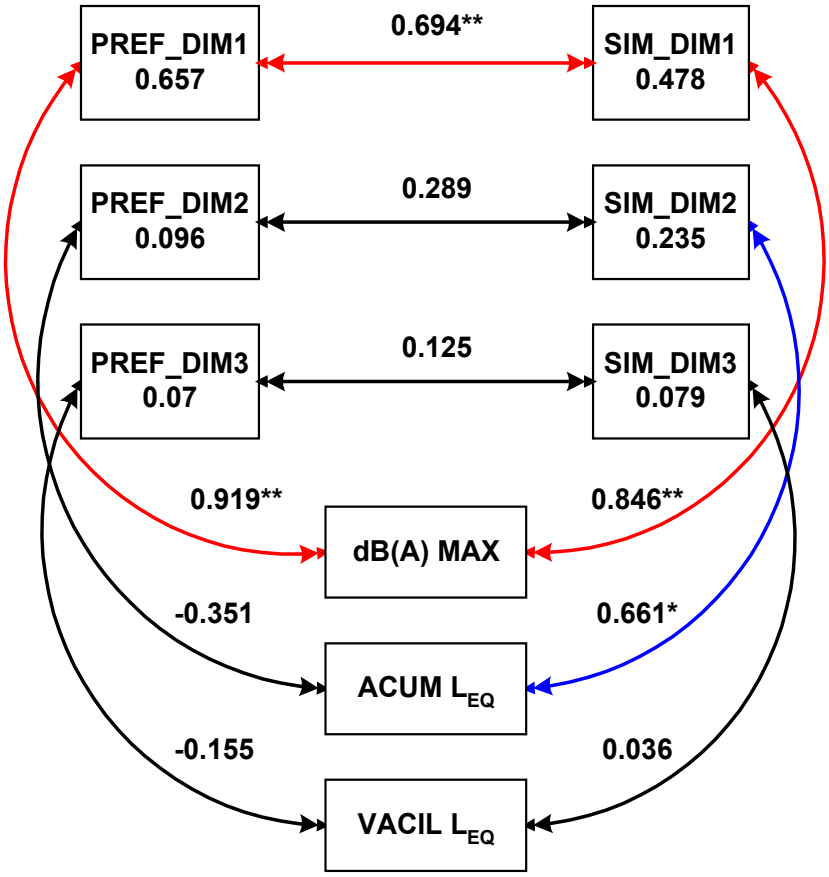

Fig. IV-10: Diagram of intercorrelations for scales obtained by two experiments and for scales measured by BAS

** correlation coefficient is significant on the level of 0,01 (two-tailed).

* correlation coefficient is significant on the level of 0,05 (two-tailed).

At least on the basis of Least Squares Regression, two physical parameters were shown to have good interrelationships with the perceptual dimensions retained from the similarity data. Technically-measurable parameters having the greatest effect on preference ratings and perceived similarity of 15 selected samples could to a large extent be described in terms of A-weighted maximum pressure level and equivalent sharpness. The results are graphically summarized in Fig. IV-10 indicating the intercorrelations between sensory attributes pertaining to preference ratings and perceived similarity and the scale values measured by the BAS. It leads us to go a step further to surmise that the perceived similarity of the same two sounds differing in dB(A) Max may either be enhanced or decreased by a change in Acum L_{eq}, and conversely. At this point, it reminds us that the resulting space with the additive constant is not Euclidean (e.g. Householder & Landahl, 1945; Attneave, 1950; Shepard, 1965). The very fact that the composite multidimensional distances X_{sum} lie approximately linear on the weighted sum of their unidimensional component distances X_{dBA} and X_{Acum}. It would seem justifiable to infer that the number of dimensions of the similarity space can be reduced by projecting perceptual dimensions into two coordinate axes, labelled dB(A) Max and Acum L_{eq} respectively, without seriously affecting the original configurations of stimulus points. The preceding discussion had imposed a restriction upon the general psychological implication with regard to the results obtained from both dimensional and nondimensional analyses applied to preference and similarity ratings. The results on the grounds of the statistical analyses of the relations between the perceptual dimensions and the signal variables are shown to be in disagreement with the choice of descriptive adjectives of sensory attributes pertaining to the judgments of preference and perceived similarity found in Section 3.3. Substantively, dimensional as well as nondimensional analyses of perceptual data indicate that two perceptual dimensions seem to underlie the ratings made by our subjects in comparing sounds. We labeled these strength and rhythm, respectively. Nevertheless, the interpretations do not appear as excessive, although our reasoning is incorrect.

Yet, there are other, often much more important sensory attributes that prompt the listener's judgments in terms of like-dislike or similar-different. The two physical parameters upon which we had focused above are not necessarily the only or the best component dimensions which to account for subjective ratings or for general usage in this research area. Consider an example. For impact sounds, such as a door slam or a hood release, or for steady state sounds, such as electric motor sounds, Kachur (1998) demonstrated that loudness consistently correlates better to subjective preference than dB(A). At this point, it might be recalled that dB(A) Max is associated with Sone Max as 0.946 and with Acum L_{eq} as 0.806, on the other hand Acum L_{eq} is associated with Asper L_{eq} as 0.703 that is also highly correlated with Vacil L_{eq} as 0.8, according to the correlation table (see Table IV-4). Those metrics may seem

interchangeable. Loudness and sharpness are related to the perception of speed variation, therefore Sone Max, Asper L_{eq}, or Vacil L_{eq} etc. are also likely to have played a decisive role in perceptual ratings, however none of these combinations could appreciably represent the similarity structure. Statistical analysis of data has shown that the similarity mechanism operates with two pronounced magnitudes, one for dB(A) Max (correlation of 85%) and the other for Acum L_{eq} (correlation of 66%), whilst the preference mechanism is dominated by a single parameter dB(A) Max (correlation of 92%). The secondary parameter Acum L_{eq} had an inferior effect on preference ratings (correlation of 35% but not significant). The following hints may be used to support this view. In a study on the relationship between 27 psychoacoustic descriptors and annoyance judgments with respect to diesel engine sounds, Khan et al. (1998) proposed a model of annoyance on the basis of loudness in sone, sharpness in acum, and harmonic ratio in dB: 92-95% of the total variance in the annoyance data was accounted for by loudness and sharpness, and 1-5% was explained by harmonic ratio that is related to body feelings. By judging the dissimilarity of vehicle noise, comparable results were reported by McAdams et al. (1998), McAdams et al. (1999), or Susini et al. (1999): A set of acoustic and psychoacoustic magnitudes had an equivalent effect on pairwise dissimilarity judgments, whereas the parameter SPL, having not being equalized, dominated pairwise preference ratings. Moreover, the view is advanced that once the sounds were equalized for SPL, other parameters came into operation for preference ratings. In this respect, it would be surmised that the perceptual dimensions extracted from the perceptual spaces are differently weighted according to the particular tasks (e.g. preference or similarity assessement) with which the subject is actually faced.

Furthermore, it is interesting to note the nongeometric, feature-matching model of psychological similarity (Restle, 1959; Tversky, 1977; Tversky & Gati, 1982) which assumes that similarity judgments are to be expressed as a linear combination of the measures of the common and distinctive features of the stimuli in that similarity increases with addition of common features and/or deletion of distinctive features. With the considerations of such a feature-mathcing process in mind, rather set-theoretically formulated than in terms of the geometric concept of distance, it is thought that both preference ratings and perceived similarity of the sounds are almost equally affected by a common feature, labelled dB(A) Max, whereas the similarity judgments are facilitated by a distinctive feature, labelled Acum L_{eq}, in sound sources under study. It is concluded that, while the two physical parameters were considered to be meaningful substantive dimensions common to all samples, the relative perceptual salience of Acum L_{eq} differed considerably depending upon the sort of judgment context, with Acum L_{eq} being much more salient in perceptual similarities than in preference ratings. At this point, Scott (1962, 1963), in reports on judgments of preference versus similarity among nations,

has pointed out that *"Simple cognitive structures tend to consist of attributes which are not well distinguished from the affective (considerations), whereas complex structures include a number of additional dimensions." (Scott, 1963, 69)*. Accordingly, it can be stated that the similarity judgments for which sharpness is a clearly correlated factor are cognitively more complex than the preference ratings. A high correlation between the patterns of the perceptual data and of the physical characteristics of the sounds would indicate that the perceptual dimensions revealed in the geometric configurations are psychologically real and might be adequate for a description of the psychoacoustic facts.

The results indicated that the perceived similarity between the pairs of sound sources can be approximated from the prediction model developed by Least Squares Regression based upon the physical parameters to serve. Seemingly, the additive model provides a better fit to the data than the Euclidean model. Keeping track of the line of thoughts from Micko and Fischer (1970, 120), the overall similarity between the pairs of sounds thus may be conceptualized either through the city-block model for the stimuli varying along two perceptually distinct attributes, which are to be represented by mutually orthogonal, independent attributes, or as an alternative by means of the non-Euclidean model for the stimuli differing with respect to two subjectively positively correlated attributes, which are probably to be represented by a rule of combination, i.e. adding. The discovery of the additivity effect and its physical basis is of importance in its own right. The upshot that two sensory attributes in a psychological space are linearly related makes it desirable to go about the additivity effect experimentally rather than statistically. As a matter of fact, the present concept attempts to delineate the processes that underlie the judgment behavior and may not be adequate for all purposes, we should admit that our additive similarities model is tentative in character since the concept of the combined "level-sharpness" effect presented so far, by which we mean the relationship between the perceptual dimensions and their physical components, is not a logically plausible relationship between the sound level and the sharpness at all. The similarity mechanism presented here can be used as a bench mark to evaluate the relevance of empirical results with different sets of sound samples.

An adequate psychophysical framework pertaining to the similarity judgments is required in that it is necessary to manipulate level or sharpness as an independent variable. The applicability of additive similarities model to describe a different sound sources is believed to be in need of further empirical verification, let alone the psychophysical study of component judgments. Going on to discuss the relationship between the auditory perception and the perceptual structure, it is expected that a change of one feature has a diagnostic meaning in order to see what effects it has on the similarity assessment. For simplicity, do louder sounds become more similar to each other if sharpness is enhanced? Or under what conditions would a

change in sound level and sharpness bring a change in the complexity of the perceptual strucutre? For the present hypothesis concering the similarity mechanism to validate "level-sharpness" interplay, both multidimensional scaling and psychophysical experimentation would be at issue. For empirical safe of extracted perceptual dimensions, McAdams et al. (1998) pointed to the BTL model ("Bradly-Terry-Luce", Bradly & Terry, 1952) with which one can establish an unidimensional order of the stimuli along the single semantic scale chosen as a basis for perceptual judgment. As part of another investigation the BTL method has been applied to the scaling of pleasantness of sounds (Ellermeier et al., 1997). Those psychophysical studies can be considered to be identification experiments faced with the perceptual problem which is prejudged as saying that certain physical orderings of the stimuli are assumed to correspond to the subject's perceptual orderings of them. The difference is that the perceptual problem is no longer prejudged in scaling experiments that inquire whether the individual stimuli can be treated as points in any multidimensional scaling space of stimulus effects (Luce, Bush, & Galanter, 1963). In reference to the points-of-view models it would be curious to know how much the perceptual structures would differ with judgment contexts, in other words, what if the subjects have to judge vehicle noises from various perspectives, as an example, from the perspective of a car seller, of a formula-one enthusiast or racer, of a noise engineer or a sound designer, of a traveller and the like. dB(A) Max and Acum L_{eq} in family-sedan might then have its virtues compared to those in sports-car. Such "what if"-scenarios would be directed toward identifying the major attributes of perspectives on sounds and thus towards what sensory attributes give rise to one or another type of perspective to be dominant.

5. Concluding Remarks

The motivation behind the study was to identify the perceptual dimensions of sounds which affect preference ratings and perceived similarity. The perceptual dimensions for which we looked for have two aspects: dimensions of the subjective equivalents retained from perceptual judgments and dimensions of the psychoacoustic correlates of sounds. We have attempted to show how the subjects determine the overall similarity of sounds from the combination of specific differences with respect to perceptually distinct stimulus features. Hence it is shown that the perceived similarity is the weighted sum of their differences resulting from two physical components on which they are represented.

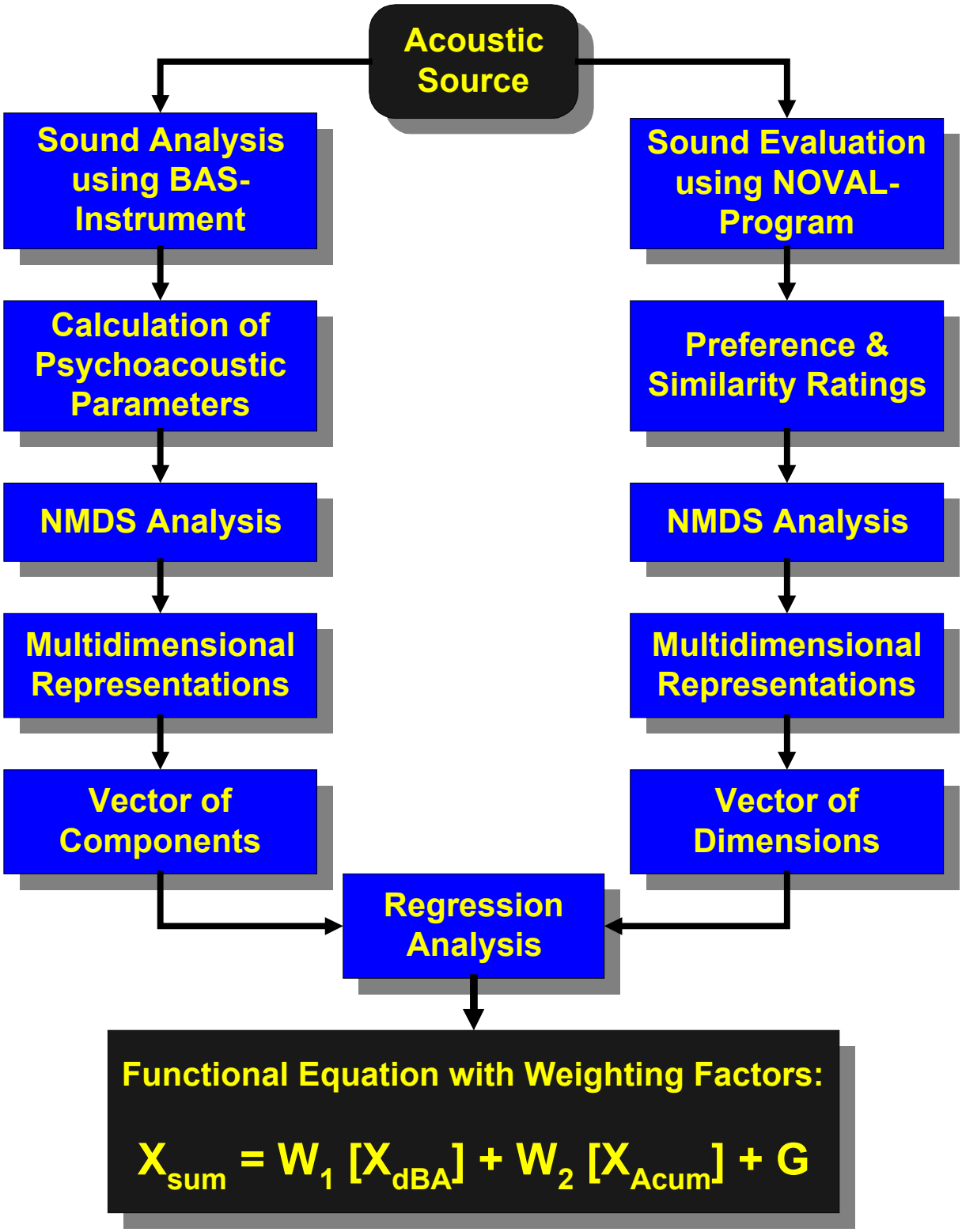

Fig. V: Model of composite rating of perceived similarity

In summary, it can be stated that nonmetric scaling has the advantage that on the one hand a structural analysis is offered to conduct phenomenological observations in a somewhat systematic fashion in order to ferret out a

parsimonious number of dimensions underlying the auditory perception. On the other hand, multivariate techniques applied to relate the perceptual dimensions to physical parameters render a simple prediction model of their interrelationship. With numerical subjective ratings of vehicle noises at hand, we looked for orderly relationships with technical data in such a way that the structural analysis based on nonmetric scaling was coupled with the statistical analysis based on Least Squares Regression. In doing so, we might be allowed to suggest that there is a composition rule on which sensory attributes of audition are to be based. The scheme, which appears in Fig. V, gives an overview of why we get a reliable MDS solution, i.e., the composite rating of perceived similarity. From the acoustic analysis of the two physical parameters, together with evaluation results, an algebraic adding model specifies how the underlying quantities are combined or integrated to give rise to overall similarity. The connection of the subjective attributes with the component dimensions would mean that functional analyses could serve as a guide for psychological dimensional research with nonmetric scaling methods, and that the opposite would be equally true. The dimensional analysis applied in both listening experiments, for the purpose of comparing the scaling spaces of preferences and similarities, clearly does not warrant for asserting that the multidimensional configurations of a set of complex sounds are indices of perceptual structures underlying the data sets. It seems worth considering here that the notion made by Torgerson, commenting *" . . . as one adds more and more obvious perceptual structure to a set of stimuli, the process underlying the similarity judgments changes from what appears to be a rather perceptual one, to one which contains more and more cognitive features. And as the contribution of cognition goes up, the appropriateness of the multidimensional representation goes down." (Torgerson, 1965, 383).*

The multidimensional scaling procedures can readily be applied to any symmetric table of data. As in Factor Analysis or Principal Component Analysis (e.g. Civille, 1998; Khan et al., 1998; for details of PCA see e.g. Martens & Naes, 1991), MDS always yields an answer, i.e, at the very least two or three dimensions can be picked up. To the detriment of MDS it is often difficult to comprehend fully the meaning of that answer. The choice of MDS is generally for two reasons: (1) MDS analysis for every subject is a fast and elegant spot-test for determining the number of psychologically meaningful dimensions to solve. (2) No a priori assumption about the physical bases of the dimensions extracted from the perceptual representations is needed. With the dimensionally organized psychololgical spaces, by means of Kruskal's algorithm, the results give an idea of how the perceptual organization of similarities can work but not exhibit firm basis for being safe of extracted perceptual dimensions as psychologically fundamental. While MDS seems to be a valuable methodological tool for analyzing perceptual structures, it would be limited in its applicability to capture the psychologically

interpretable dimensions especially if the dimensionality of the Minkowski spaces exceeds two. Suffice it to say that we can get a spatial representation of proximity data, and that the criterion which to decide against upon the best-fit solution of points and the proper value of dimensionality is the sense that can be made of it, and that the outcome can be retrieved or predicted from it. Accompanied by MDS, Least Squares Regression proves to be a viable method for a more reliable interpretation about the nature of the perceptual dimensions. Thus the combination of MDS and multivariate techniques contribute to the understanding of the MDS solutions underlying the perceptual representations. To capture the proper number of dimensions in a dimensionally organized space in a parsimonious way and to properly infer the nature of the stimulus objects, MDS coupled with Least Squares Regression has to be applied for the structural analysis of any similarity data.
So far, we had showed that MDS could serve as a useful tool for determinining a parsimonious number of dimensions required to represent the psychologically meaningful attributes of sounds that underlie the perceptual judgments. The studies presently described offer a relatively small contribution to an understanding of the nature of the multidimensional configurations of points based on preference ratings and similarity judgments. The mental model of sound perception which we have touched is constructed at large by interpreting the perceived similarity and its multidimensional structure. As the relationships between technically-measurable parameters and their audible importance become clearer, it might be possible to design "better" product sound and so to develop a automotive product with "winner" sound quality in the marketplace. Although smaller in research scale, it is to be hoped that our results concerning the "level-sharpness" effect represents a major portion of perceptual similarity which might offer positive challenges for further work in this area.

Summary

A series of experiments that touched a variety of nonmetric multidimensional scaling of preference and similarity relations on a set of vehicle interior noises was performed, gathered from a small group of subjects from the University of Oldenburg, by means of the paired comparison method. The principal objective of this study was to discover the best-ftting configuration for the 15 different vehicle interior noises in an extended Euclidean minimum dimensionality space such as the ordering of pairs of samples concerning their dissimilarities, which are tied in with those relative to their transformed metric distances. Experiment I was primarily concerned with applying Kruskal's algorithm to pairwise preference ratings, and Experiment II to

pairwise similarity judgments. The Euclidean model compared with rivals could quite well account for structural properties of subjective noise assessments. Since the same samples were used in both experiments, the comparison of the multidimensional representations was possible. Similarity evaluation might have been influenced by an additional attribute that appeared to have only a slight effect on preference evaluation. Ss just needed two dimensions for perceived similarity, and the perceptual dimensions could be linked to the physical parameters identified. Least Squares Regression of similarity data attained in the results quite favor an algebraic adding model. The principal acoustic factors for perceived similarity with regard to a set of vehicle interior noise appeared as the weighted sum of dB(A) Max and Acum L_{eq}, i.e. the perceptual ratings are weighted and summed by two key attributes – sound level and sharpness. Summarizing the findings obtained from the structural and the analytical analysis of perceptual data, it can be concluded:

(1) The MDS solutions are found to fit quite well to a 2D Euclidean model for both preference ratings and similarity judgments.

(2) The principal axes in the multidimensional representations with regard to preference ratings and similarity judgments are related to a common judgment criterion. Preference structure is shown to be a subset of similarity structure.

(3) Preference ratings vary only along dB(A) Max, whilst similarity judgments mainly comprise levels in dB(A) Max and sharpness in Acum L_{eq}.

(4) Preference ratings show a gaining effect of 1.66 dB(A) Max, whilst similarity judgments estimate about 2 dB(A) Max greater than the technically measured quantities of sound sources.

(5) Perceived similarity is reducible to a matter of calculating the combination of two physical components, dB(A) Max and Acum L_{eq}, each of which has differential weighting, i.e. perceived similarity is approximately proportional to the weighted sum of dB(A) Max and Acum L_{eq}.

(6) MDS provides an elegant and pragmatical solution for a parsimonious number of dimensions in the data, yet it cannot guarantee that the principal axes of the dimensionally organized space are considered to be psychologically meaningful dimensions common to all sound samples in interest.

(7) The application of MDS coupled with Least Squares Regression contributes to a reliable interpretation about psychological dimensions.

References

Abeslson, R. P. (1955). A technique and a model for multidimensional attitude scaling. Public Opinion Quaterly, Vol. 18, 405-418.

Abelson, R. P. (1963). Computer simulation of “hot cognitions“. In Tomkins, S., & Mesick, S. (Eds.), Computer simulation of personality. New York: Wiley.

Abelson, R. P., & Sermat, V. (1962). Multdimensional scaling of facial expressions. Journal of Experimental Psychology, Vol. 63, No. 6, 546-554.

Ahrens, H. J. (1970). Zur Verwendung des Metrik-Parameters multidimensionaler Skalierungen bei der Analyse von Wahrnehmungsstrukturen. Zeitschrift für Experimentelle und Angewandte Psychologie, Band XIX/2, 173-195.

Ahrens, H. J. (1974). Multidimensionale Skalierung. Weinheim: Beltz.

Albersheim, G. (1975). Zur Psychologie der Toneigenschaften unter Beruecksichtigung der “Zweikomponenten-Theorie“ und der Vokalsystematik. Verlag Valentin Koerner. Baden-Baden.

Anderson, A. B. (1970). Brief report: the effect of aggregation on nonmetric multidimensional scaling solutions. Multivariate Behavioral Research, Vol. 5, 369-373.

Anderson, A. B. (1973). Brief report: Additional data on distortion due to aggregation in nonmetric multidimensional scaling. Multivariate Behavioral Research, 8/3, 519-523.

Anderson, N. H. (1973a). Algebraic models in perception. In Carterette, E. C., & Friedman, M. P. (Eds.), Handbook of perception. New York: Academic Press, 1973.

Anderson, N. H. (1973b). Cognition algebra: Integration theory applied to social attribution. In Berkowitz, L. (Ed.), Advances in experimental social psychology. Vol. 7. New York: Academic Press, 1973.

Anderson, N. H. (1973c). Functional measurement of social desirability. Sociometry, Vol. 36, No. 1, 89-98.

Anderson, N. H. (1974). Information integration theory: A brief survey. In Krantz, D. H., Atkinson, R. C., Luce, R. D., & Suppes, P. (Eds.), Measurement, psychophysics, and neural information processing. Vol. II. San Francisco: Freeman, 1974.

ANSI S1.11 (1986). American national standard for specification for octave-band and fractional-octave-band analog and digital filters.

Arabie, P. (1991). Was Euclid an unnecessarily sophisticated psychologist? Psychometrika, 56, 567-587.

Attneave, F. (1950). Dimensions of similarity. Am. J. Psych., Vol. 63, 516-556.

Aures, W. (1984a). Berechnungsverfahren für den Wohlklang beliebiger Schallsignale, ein Beitrag zur gehörbezogenen Schallanalyse. Unpublished doctoral dissertation. Technical University of München.

Aures, W. (1984b). Ein Berechnungsverfahren der Rauhigkeit. Acustica, Vol. 58, 268-281.

Aures, W. (1984c). Der sensorische Wohlklang als Funktion psychoakustischer Empfindungsgrößen. Acustica, Vol. 58, 282-290.

Aures, W. (1985). Berechnungsverfahren für den sensorischen Wohlklang beliebiger Schallsignale. Acustica, Vol. 59, 130-141.

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (1994). Multivariate Analysemethoden. Eine anwendungsorientierte Einführung. Springer-Verlag.

Bäumer, H.-P. (2000). Räumliche Statistik. Anwendungen in Biologie und Ökologie. Unpublished manuscript. University of Oldenburg.

Beals, R., & Krantz, D. H. (1967). Metrics and geodesics induced by order relations. Mathematische Zeitschriften, Vol. 101, 285-298.

Beals, R., Krantz, D. H., & Tversky, A. (1968). Foundations of multidimensional scaling. Psychological Review, Vol. 75, No. 2, 127-142.

Beidl, C. & Stücklschwaiger, W. (1997). Application of the AVL-annoyance index for engine noise quality development. Acustica, Vol. 83, 789-795.

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

Berglund, U. (1981). Loudness and annoyance from comunity noises. In: Schick, A. (Ed.). Akustik zwischen Physik and Psychologie. Klett-Cotta, Stuttgart, 27-32.

Berlyne, D. E. (1966). Curiostiy and exploration. Science, Vol. 153, 23-33.

Berlyne, D. E. (1975). Dimensions of perception of exotic and pre-renaissance paintings. Canadian Journal of Psychology, Vol. 29, No. 2, 151-173.

Berlyne, D. E., & Ogilvie, J. C. (1974). Dimensions of perception of paintings. In Berlyne, D. E. (Ed.), Studies in the new experimental aesthetics. Washington: Hemisphere.

Birnbaum, M. H. (1973). The Devil rides again: Correlation as an index of fit. Psychological Bulletin, Vol. 79, 239-242.

Bismarck, G. v. (1972). Extraktion und Messung von Merkmalen der Klangfarbenwahrnehmung stationärer Schalle. Unpublished doctoral dissertation. Technical University of München.

Bismarck, G. v. (1974a). Timbre of steady sounds: A factorial investigation of its verbal attributes. Acustica, Vol. 30, 146-159.

Bismarck, G. v. (1974b). Sharpness as an attribute of the timbre of steady sounds. Acustica, Vol. 30, 159-172.

Bisping, R. (1994). Psychologische und akustische Aspekte der Klangqualität von Fahrzeuggeräuschen. In Vo, Q. H. (Ed.). Sound engineering. Expert Verlag, Renningen-Malmsheim, 34-45.

Bisping, R. (1995). Emotional effect of car interior sounds: Pleasantness and power and their relation to acoustic key features. Noise and Vibration Conference 1995, Vol. 2, SAE., I. Society of Automotive Engineers (Ed.). Warrendale, 1203-1209.

Bisping, R. (1997). Car interior sound quality: Experimental analysis by synthesis. Acustica, Vol. 83, 813-818.

Bisping, R. & Giehl, S. (1996). Psychological analysis of the sound quality of vehicle interior noise: Field and laboratory experiments. Proceedings of the 1996 Engine and Environment Conference. AVL List GmbH, Graz, 79-86.

Blauert, J. (1986). Cognitive and aesthetic aspects of noise engineering. In: Proc. Inter-noise ´86, Vol., I., 5-14.

Blauert, J. (1996). “Sound Quality”. EAA Tutorium, Antwerp.

Blauert, J., & Jekosch, U. (1997). Sound quality evaluation. A multi-layered problem. Acustica, Vol. 83, 747-753.

Bodden, M. (1997). Instrumentation for Sound Quality Evaluation. Acustica, 83, 775-783.

Bodden, M. (2000). Allgemeine Vorgehensweise Sound Quality. In: Bodden, M., & Heinrichs, M. (Eds.). Geräuschqualität: Methoden und Umsetzung in der Fahrzeug- und Elektroindustrie. Tagung Nr. H030-06-012-0. Haus der Technik, Essen.

Boemak, N. (1994). Bewertung von Motorgeräuschen durch subjektiven Vergleich im Akustiklabor. Zeitschrift für Lärmbekämpfung, Vol. 41, 84-88.

Borg, I., & Groenen, P. (1997). Modern multidimensional scaling: Theory and applications. New York: Springer.

Borg, I., & Lingoes, J. (1987). Multdimensional similarity structure analysis. Springer.

Bortz, J. (1974). Kritische Bemerkungen über den Einsatz nichteuklidischer Metriken im Rahmen der multidimensionalen Skalierung. Archiv für Psychologie, 126, 194-212.

Bortz, J. (1993). Statistik für Sozialwissenschaftler. Springer-Verlag.

Bradly, R. A., & Terry, M. E. (1952). The rank analysis of incomplete block designs. I. The method of paired comparisons. Biometrika, 39.

Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland forest communications of Southern Wisconsin. Ecological Monographs, Vol. 27, 325.

Brennecke, W., & Remmers, H. (1983). Physikalische Parameter bei der Bewertung der Lästigkeit von Industriegeräuschen. Acustica, Vol. 52, 279-289.

Brunswik, E., & Kamiya, J. (1953). Ecological cue-validity of proximity and of other Gestalt factors. Amer. J. Psychol., Vol. 66, 20-32.

Cardozo, B. L., & Van der Veen, K. G. (1979). Estimation of annoyance due to low level sound. Applied Acoustics, Vol. 12, 389-396.

Cardozo, B. L., & Van Lieshout, R. A. J. M. (1981a). Estimates of annoyance of sounds of different character. Applied Acoustics, Vol. 14, 323-329.

Cardozo, B. L., & Van Lieshout, R. A. J. M. (1981b). Splitting sound annoyance into two components: level and sound character. IPO Annual Progress Report, Vol. 16, 23.

Carnap, R. (1967). The logical structure of the world. Pseudoproblems in philosophy. London: Routledge and Kegan Paul.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of Echart-Young's decomposion. Psychometrika, Vol. 35, 283-319.

Carroll, J. D., & Wish, M. (1974a). Multidimensional perceptual models and measurement methods. In Carterette, E., & Friedman, M. (Eds.), Handbook of perception. Vol. II, New York: Academic Press.

Carroll, J. D., & Wish, M. (1974b). Models and methods for three-way multidimensional scaling. In Krantz, D., Atkinson, R., Luce, R., & Suppes, P. (Eds.), Measurement, psychophysics, and neural information processing. Vol. II. San Francisco: Freeman.

Choe, B. (2000a). Multidimensional scaling of car interior noise: structural analysis of acoustical preference judgment. Proceedings of the first KRISS International Autumn Session in Ergonomics, Taejon.

Choe, B. (2000b). A nonmetric multidimensional analysis of 15 complex sounds: a comparison of preferences and similarities evaluations. Reports from the Institute for Research into Man-Environment-Relations, No. 35, November 2000, Oldenburg.

Chouard, N. (1997). Loudness and unpleasantness perception in dichotic conditions. Unpublished doctoral dissertation. University of Oldenburg.

Civille, G. V. (1998). Understanding the flavors of foods and beverages in terms of subjective and objective sensory evaluation responses. In: Davies, P., & Ebbitt, G. (Eds.). Sound Quality Symposium, Ypsilanti Michigan USA, 17-22.

Cliff, N. (1968). The "idealized individual" interpretation of individual differences in multidimensional scaling. Psychometrika, Vol. 33, 225-232.

Cliff, N. (1969). Liking judgments and multidimensional scaling. Educational and Psychological Measurement, Vol. 29, 73-86.

Cliff, N., Bradley, P., & Girard, R. (1973). The investigation of cognitive models for inventory response. Multivariate Behavioral Research, Vol. 8, 407-426.

Cliff, N., & Young, F. W. (1968). On the relation between the unidimensional judgments and multidimensional scaling. Organizational Behavior and Human Performance, 3, 269-285.

Coombs, C. H. (1950). Psychological scaling without a unit of measurement. Psychological Review, Vol. 57, 145-158.

Coombs, C. H. (1951). Mathematical models in psychological scaling. Journal of the American Statistical Association, Vol. 46, 480-489.

Coombs, C. H. (1958). An application of a nonmetric model for multidimensional analysis of similarities. Psychol. Rep., Vol. 4, 511-518.

Coombs, C. H. (1960). A theory of data. Psychological Review, Vol. 67, No. 3, 143-159.

Cooper, L. G. (1973). A multivariate investigation of preferences. Multivariate Behavioral Research, Vol. 8, 253-272.

Cox, M. A. A., & Cox, T. F. (1992). Interpretation of Stress in non-metric multidimensional scaling. Statistica Applicata, Vol. 4, 611-618.

Critchley, F., & Fichet, B. (1994). The partial order by inclusion of the principal classes of dissimilarity on a finite set and some of their basic properties. In Cutsem, B. V. (Ed.), Classification and dissimilarity analysis. Vol. 93. New York: Springer. pp. 5-65.

Daniel, P. (1997). Berechnung und kategoriale Beurteilung der Rauhigkeit und Unangenehmheit von synthetischen und technischen Schallen. Published doctoral dissertation. University of Oldenburg. Fortschritt-Berichte VDI.

Daniel, P., & Weber, R. (1993). Calculating psychoacoustical roughness. In Schick, A. (Ed.). Contributions to psychological acoustics. Results of the sixth oldenburg symposium on psychological acoustics. bis, Oldenburg, 251-265.

David, H. A. (1988). The method of paired comparisons. Oxford University Press.

de la Motte-Haber, H. (1972). Musikpsychologie. Eine Einfuehrung. Musikverlag. Hans Gerig Koeln.

Diederich, G., Messick, S. J., & Tucker, L. R. (1957). A general least squares solution for successive intervals. Psychometrika, Vol. 22, 159-173.

DIN 45 631 (1967). Berechnung des Lautstärkepegels aus dem Geräuschspektrum. Verfahren nach E. Zwicker.

DIN 45 639 (1969). Innengeräuschmessungen in Kraftfahrzeugen (inside noise of motor vehicles).

DIN 45 681 (1992). Bestimmung der Tonhaltigkeit von Geräuschen und Ermittlung eines Tonzuschlages für die Beurteilung von Geräuschimmissionen (Detection of tonal components of noise and determination of a tone adjustment for the assessment of noise immissions).

Donnadieu, S., & McAdams, S. (1996). Effects of context on dissimilarity, discrimination and categorisation task on timbre perception. Proceedings of the 12. Annual Meeting of the International Society for Psychophysics, Fechner Day 96, Padua, 239-244.

Eckart, C., & Young, F. W. (1936). The approximation of one matrix by another of lower rank. Psychometrika, Vol. 1, 211-218.

Edwards, A. L. (1952). The scaling of stimuli by the method of successive intervals. J. Appl. Psychol., 36, 118-122.

Eisler, H., & Ekman, G. (1959). A mechanism of subjective similarity. Acta Psychologica, Vol. 16, 1-10.

Ekman, G. (1954). Dimensions of color vision. Journal of Psychology, Vol. 38, 467-474.

Ekman, G. (1965). Two methods for the analysis of perceptual dimensionality. Perceptual and Motor Skills, Vol. 20, 557-572.

Ekman, G., Engen, T., Künnapas, T., & Lindman, R. (1964). A quantitative principle of qualitative similarity. Journal of Experimental Psychology, Vol. 68, 530-536.

Ellermeier, W., Mader, M. & Daniel, P. (1997). Scaling the unpleasantness of natural sounds according to the BTL model. Fortschritte der Akustik, DAGA 97, 407-408. Oldenburg: DEGA.

Fahrmeir, L. & Hamerle, A. (1984). Multivariate statistische Verfahren. Berlin etc.: de Gruyter.

Fastl, H. (1982). Fluctuation strength and temporal masking patterns of amplitude modulated broadband noise. Hearing Research, 8, 59-69.

Fastl, H. (1983). Fluctuation strength of modulated tones and broadband noise. In: Klinke, R., & Hartmann, R. (Eds.). Hearing – physiological bases and psychophysics. Springer, Berlin, 282-288.

Fastl, H. (1991). Beurteilung und Messung der wahgenommenen äquivalenten Dauerlautheit. Zeitschrift für Lärmbekämpfung, Vol. 38.

Fastl, H. (1997). The psychoacoustics of sound-quality evaluation. Acustica, 83, 754-764.

Fastl, H. (1998). Psychoacoustics and sound quality metrics. In: Davies, P., & Ebbitt, G. (Eds.). Sound Quality Symposium, Ypsilanti Michigan USA, 3-10.

Fricke, J., Moser, L. M., Scheurer, H., & Schubert, G. (1983). Schall und Schallschutz. Grundlagen und Anwendungen. Physik-Verlag. Weinheim.

Gabrielsson, A., & Lindström, B. (1985). Perceived sound quality of high-fidelity loudspeakers. J. Audio Eng. Soc., Vol. 33, No. 1/2, 33-53.

Garner, W. R. (1962). Uncertainty and structure as psychological concepts. New York: Wiley.

Garner, W. R. (1974). The processing of information and structure. Potomac, MD: Erlbaum.

Gati, I., & Tversky, A. (1982). Representations of qualitative and quantitative dimensions. Journal of Experimental Psychology: Human Perception and Performance, 8, 325-340.

Genuit, K. (1997). Background and practical examples of sound design. Acustica, Vol. 83, 805-812.

Genuit, K. (2000). SoundCar. Verfahren zur kombinierten Beurteilung von Akustik und Schwingungen im Kfz-Innenraum. In: Geräuschqualität: Methoden und Umsetzung in der Fahrzeug- und Elektroindustrie. Tagung Nr. H030-06-012-0. Haus der Technik, Essen.

Gigerenzer, G. (1977). Nonmetrische multidimensionale Skalierung als Modell des Urteilsverhaltens. Zur Integration von dimensionsanalytischer Methodik und psychologischer Theoriebildung. Unpublished doctoral dissertation. University of München.

Gigerenzer, G. (1981). Messung und Modellbildung in der Psychologie. Ernst Reinhardt Verlag München Basel.

Goodall, D. W. (1966). A new similarity index based on probability. Biometrics, Vol. 2, 882-907.

Gower, J. C. (1966). Some distance properties of latent root and vector methods using multivariate analysis. Biometrika, Vol. 53, 325-338.

Graef, J., & Spence, I. (1979). Using distance information in the design of large multidimensional scaling experiments. Psychological Bulletin, Vol. 86, No. 1, 60-66.

Green, P. E., & Carmone, F. J. (1969). Multidimensional scaling: An introduction and comparison of nonmetric unfolding techniques. J. Mark. Res., Vol. 6, 330-341.

Green, P. E., Maheshwari, A., & Rao, V. R. (1969). Dimensional interpretation and configuration invariance in multidimensional scaling: an empirical study. Multivariate Behavioral Research, Vol. 4, 159-180.

Gregson, R. A. M. (1975). Psychometrics of Similarity. Academic Press: New York.

Gregson, R. A. M. (1976). A comparative evaluation of seven similarity models. Br. J. math. Statist. Psychol., Vol. 29, 139-156.

Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. J. Acoust. Soc. Am., Vol. 61, No. 5, 1270-1277.

Guski, R. (1997). Psychological methods for evaluating sound quality and assessing acoustic information. Acustica, Vol. 83, 765-774.

Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. Psychometrika, Vol. 33, 469-506.

Guttman, L., & Kalish, H. I. (1956). Discriminability and stimulus generalization. Journal of Experimental Psychology, Vol. 51, 79-88.

Harrison, A. A. (1968). A response competition, frequency, exploratory behavior, and liking. Journal of Personality and Social Psychology, Vol. 9, 363-368.

HEAD acoustics GmbH (1996). Binaural Analysis System Manual. Herzogenrath.

Heldmann, K. (1994). Wahrnehmung, gehörgerechte Analyse und Merkmalsextraktion technischer Schalle. Unpublished doctoral dissertation. Technical University of München.

Hellbrück, J. (1986). Psychologische Beiträge zur Audiologie und Audiometrie. Struktur und Dynamik des Hörens. Habilitationsschrift. University of Würzburg.

Hellbrück, J. (1991). Loudness Scaling: Ratio Scales, Rating Scales, Intersubject Variability and some other Problems. In Schick, A. (Hrsg.). Contributions to Psychological Acoustics: Results of the 5th Oldenburg Symposion on Psychological Acoustics, 119-133.

Heller, O. (1985). Hörfeldaudiometrie mit dem Verfahren der Kategorienunterteilung (KU). Psychologische Beiträge, Vol. 27, 478-493.

Helm, C. E., & Tucker, L. R. (1962). Individual differences in the structure of color-perception. American Journal of Psychology, Vol. 75, 437-444.

Helmholtz, H. (1863). Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik. Braunschweig: Vieweg und Sohn.

Horan, C. B. (1969). Multidimensional scaling: Combining observations when individuals have different perceptual structures. Psychometrika, Vol. 34, 139-165.

Hornbostel, E. M. (1926). Psychologie der Gehörerscheinungen. In: Bethe, A. (Hrsg.): Handbuch der normalen und pathologischen Physiologie. Berlin, 701-730.

Householder, A. S., & Landahl, H. D. (1945). Mathematical biophysics of the central nervous system. Bloomington, Indiana: Principia Press.

Howard, J. H., & Silverman, E. B. (1976). A multidimensional scaling analysis of 16 complex sounds. Perception & Psychophysics, Vol. 19, 193-200.

Hyman, R. & Well, A. (1968). Perceptual separability and spatial models. Perception and Psychophysics, Vol. 3, 161-165.

Isaac, P. D. (1970). Dissimilarities as indices of individual perceptual structure. Perception & Psychophysics, Vol. 7, 229-233.

ISO 5128 (1980). Measurement of noise inside motor vehicles.

ISO 532 (1975). Acoustics – Method for calculating loudness level.

ISO/R 362 (1964). Measurement of noise emitted by motor vehicles.

Johnson, S. C. (1967). Hierarchical clustering schemes. Psychometrika, 32/3, 241-254.

Jost, E. (1967). Akustische und psychometrische Untersuchungen an Klarinettenklaengen. Veroeffentlichungen des Staatlichen Institus fuer Musikforschung. Bd. 1, Koeln.

Jost, E. (1973). Über Klangeigenschaften von Lautsprechern. Eine experimentalpsychologische Untersuchung, Jahrbuch des Staatlichen Institutes für Musikforschung, Berlin.

Kachur, M. (1998). A survey of sound quality jury evaluation correlations: loudness versus A-weighted sound level. In: Davies, P., & Ebbitt, G. (Eds.). Sound Quality Symposium, Ypsilanti Michigan USA, 71-76.

Khan, M. S., Johansson, Ö., & Sundbäck, U. (1998). A laboratory study of sound quality with artificial head and stereophonic recordings in free field conditions. In: Davies, P., & Ebbitt, G. (Eds.). Sound Quality Symposium, Ypsilanti Michigan USA, 43-52.

Kastka, J. (1981). Psychologische Indikatoren der Verkehrslärmbelästigung. In: Schick, A. (Ed.). Akustik zwischen Physik und Psychologie. Klett-Cotta, Stuttgart, 68-86.

Keiper, W. (1997). Sound quality evaluation in the product cycle. Acustica, 83, 784-788.

Kendall, R. A., & Carterette, E. C. (1991). Perceptual scaling of simultaneous wind instrument timbres. Music Perception, Vol. 8, 369-404.

Klahr, D. (1969). A Monte Carlo investigation of the statistical significance of Kruskal's nonmetric scaling procedure. Psychometrika, Vol. 34, No. 3, 319-330.

Koehler, W. (1923). Tonpsychologie. Handbuch der Neurologie des Ohres. Alexander & Marburg (Hrsg.). Berlin und Wien.

Kohler, M. & Kotterba, B. (1992). Die Fusion gemessener Empfindungsgrößen zur objektiven Bestimmung des Wohlklanges. In Fortschritte der Akustik, DAGA, 905-908.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). Foundations of measurement: additive and polynomial representations. Vol. I. New York: Academic Press.

Krantz, D. H., & Tversky, A. (1975). Similarity of rectangles: An analysis of subjective dimensions. Journal of Mathematical Psychology, Vol. 12, 4-34.

Krumhansl, C. L. (1989). Why is musical timbre so hard to understand? In Nielzén, S., & Olsson, O. (Eds.), Structure and perception of electroacoustic sound and music. Amsterdam: Elsevier, 43-53.

Kruskal, J. B. (1964a). Multdimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika, Vol. 29, No. 2, 1-27.

Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. Psychometrika, Vol. 29, No. 2, 115-129.

Kryter, K. D. (1967). Concepts of perceived noisiness, their implementation and application. J. Acoust. Soc. Am., Vol. 43, No. 2, 344-361.

Künnapas, T. (1966). Visual perception of capital letters: multidimensional ratio scaling and multidimensional similarity. Scandinavian Journal of Psychology, Vol. 7, 189-196.

Landahl, H. D. (1945). Neural mechanism for the concepts of difference and similarity. Bulletin of Mathematical Biophysics, Vol. 7, 83-88.

Laucken, U., Mees, U., & Schick, A. (1987). The position of annoyance within the logographic structure of letters of complaint. In: Koelega, H. S. (Ed.). Environmental annoyance. Characterization, measurement, and control. Amsterdam: Elsevier, 383-397.

Laux, P. C., & Davies, P. (1998). Procedure to develop an artificial neural network model that predicts annoyance to time-varying noises. In: Davies, P., & Ebbitt, G. (Eds.). Sound Quality Symposium, Ypsilanti Michigan USA, 85-90.

Lazarte, A., & Schönemann, P. H. (1991). Saliency metric for subadditive dissimilarity judgments of rectangles. Perception and Psychophysics, 49, 142-158.

Levelt, W. J. M., Van de Geer, J. P., & Plomp, R. (1966). Triadic comparisons of musical intervals. British Journal of Mathematical and Statistical Psychology, Vol. 19, 163-179.

Lingoes, J. C. (1971). Some boundary conditions for a monotone analysis of symmetric matrices. Psychometrika, Vol. 36, 195-203.

Lingoes, J. C. & Roskam, E. E. (1973). A mathematical and empirical analysis of two multidimensional scaling algorithms. Psychometrika Monograph Supplement, Vol. 38.

Luce, R. D., Bush, R. R., & Galanter, E. (1963). Handbook of mathematical psychology. Vol. I, New York: Wiley.

Martens, H. & Naes, T. (1991). Multivariate calibration. Wiley, New York.

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimhoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. Psychological Research, Vol. 58, 177-192.

McAdams, S., Susini, P., Misdariis, N., & Winsberg, S. (1998). Multidimensional characterisation of perceptual and preference judgments of vehicle and environmental noises. Euro-Noise 98: Designing for Silence, Muenchen, 561-566.

McAdams, S., & Winsberg, S. (1999). Multidimensional scaling of musical timbre contrained by physical parameters. Journal of the Acoustical Society of America, Vol. 105, 1273(A).

McAdams, S., & Winsberg, S. (2000). Psychophysical quantification of individual differences in timbre perception. In: Schick et al. (Eds.). Contributions to Psychological Acoustics: Results of the 8th Oldenburg Symposium on Psychological Acoustics. Oldenburg: bis, v.

McCulloch, W. S., & Pitts, W. (1947). How we know universals: The perception of auditory and visual forms, Bull. Math. Biophysics, Vol. 9, 127-148.

McGee, V. E. (1968). Multidimensional scaling of n sets of similarity measures: a nonmetric individual differences approach. Multivariate Behavioral Research, 3, 233-248.

McQuitty, L. L. (1960). Hierarchical linkage analysis for the isolation of types. Educational and Psychological Measurement, Vol. 20, 55-67.

Mellert, V. & Weber, R. (1981). Physikalische Faktoren der Lästigkeit. In: Schick, A. (Ed.): Akustik zwischen Physik und Psychologie. Klett-Cotta, Stuttgart, 48-62.

Messick, S. J. (1956). An empirical evaluation of multidimensional successive intervals. Psychometrika, Vol. 21, 367-376.

Micko, H. C., & Fischer, W. (1970). The metric of multidimensional psychological spaces as a function of the differential attention to subjective attributes. Journal of Mathematical Psychology, Vol. 7, 118-143.

Miller, G. A. (1969). A psychological method to investigate verbal concepts. Journal of Mathematical Psychology, Vol. 6, 169-191.

Miller, J. R., & Carterette, E. C. (1975). Perceptual space for musical structures. J. Acoust. Soc. Am., Vol. 58, No. 3, 711-720.

Milord, J. T. (1978). Aesthetics aspects of faces: A (somewhat) phenomenologic analysis using multidimensional scaling methods. Vol. 36, No. 2, 205-216.

Minkoswki, H. (1910). Geometrie der Zahlen. Leipzig. New York: Johnson. 1968.

Misdariis, N., Smith, B., Pressnitzer, D., Susini, P., McAdams, S. (1998). Validation of a multidimensional distance model for perceptual dissimilarities among musical timbres. J. Acoust. Soc. Am. 103, 3005.

Mueller, W. S. (1974). Cognitive complexity and salience of dimensions in person perception. Australian Journal of Psychology, Vol. 26, 173-182.

Nelson, L. D., & Nelson, D. L. (1970). Rated acoustic (articulatory) similarity for word pairs varying in number and ordinal position of common letters. Psychonomic Science, Vol. 19, 81-82.

O'Hare, D. (1976). Individual differences in perceived similarity and preference for visual art: A multidimensional scaling analysis. Perception & Psychophysics, 20(6), 445-452.

Olson, R. K., & Attneave, F. (1970). What variables produce similarity grouping? American Journal of Psychology, Vol. 83, 1-21.

Osgood, C. E. (1949). Meaningful similarity and interference in learning. Journal of Experimental Psychology, Vol. 36, 277-301.

Osgood, C. E. (1962). Studies on the generality of affective meaning systems. American Psychologist. Vol. 17, 10-28.

Plomp, R. (1970). Timbre as a multidimensional attribute of complex tones. In Plomp, R. & Smoorenburg, G. F. (Eds.), Frequency analysis and periodicity detection in hearing. Leiden: Sijthoff, 397-414.

Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth.

Plomp, R., Pols, L. C. W. & van der Geer, J. P. (1965). Dimensional analysis of vowel spectra. J. Acoust. Soc. Am. 38, 548-560.

Pols, L. C. W., van der Kamp, J. T. & Plomp, R. (1969). Perceptual and physical space of vowel sounds. J. Acoust. Soc. Am. 46, 458-467.

Preis, A. (1987). Intrusive Sounds. Applied Acoustics, Vol. 20, 101-127.

Restle, F. (1959). A metric and an ordering on sets. Psychometrika, Vol. 24, 207-220.

Richardson, M. W. (1938). Multdimensional psychophysics. Psych. Bull., 35, 659-660.

Russell, B. (1940). An inquiry into meaning and truth. George Allen and Unwin Ltd.

Russell, P. N., & Gregson, R. A. (1967). A comparison of intermodal and intramodal methods in multidimensional scaling of three-component taste mixtures. Australian Journal of Psychology, Vol. 18, 244-254.

Sattath, S., & Tversky, A. (1977). Additive similarity trees. Psychometrika, 42, 319-345.

Schick, A. (1979). Schallwirkung aus psychologischer Sicht. Stuttgart: Klett-Cotta

Schick, A. (1990a). On the validity of annoyance judgments: An overview on different viewpoints. In: Müller, Fr. (Ed.): Proceedings of the Sixth Annual Meeting of the International Society for Psychophysics. Würzburg 1990, F.R.G. Fechner Day 90. Würzburg: International Society for Psychophysics, 146-151.

Schick, A. (1990b). Schallbewertung. Grundlagen der Lärmbewertung. Springer-Verlag.

Schick, A. (1992). Aktuelle Probleme der Schallwirkungsforschung. HNO, 40, Springer-Verlag, 37-40.

Schick, A. (1994). Zur Geschichte der Bewertung von Innengeraeuschen in Personenwagen. Zeitschrift fuer Lärmbekämpfung, 41, Springer-Verlag, 61-68.

Schick, A. (1995). Geräusche im Fahrgastraum des PKW: ein geschichtlicher Abriß. In: Ahmed, S. R. (Ed.). Akustik und Aerodynamik des Kraftfahrzeuges. Expert Verlag. 32-47.

Schick, A. (1996). The role of subjects and its importance for differential noise psychology. In: Inter'Noise 96. Proceedings. The 1996 International Congress on Noise Control Engineering. Liverpool-England, Vol. 4, 2121-2126.

Schick, A. (2000). Semantics of hearing in car manufacturing. Seventh International Congress on Sound and Vibration. Garmisch-Partenkirchen, Germany, 2519-2526.

Schiffman, S., & Dackis, C. (1975). Taste of nutrients: Amino acids, vitamins, and fatty acids. Perception & Psychophysics, Vol. 17, 140-146.

Schlosberg, H. (1954). Three dimensions of emotion. Psychological Reveiw, 61, 81-88.

Schole, H. (1930). Tonpsychologie und Musikaesthetik. Goettingen.

Schönemann, P. H. (1994). Measurement: The reasonable ineffectiveness of mathematics in the social sciences. In Borg, I., & Mohler, P. P. (Eds.), Trends and perspectives in empirical social research. New York: De Gruyter. 149-160.

Schönemann, P. H., & Borg, I. (1983). Grundlagen der metrischen mehrdimensionalen Skaliermethoden. In Feger, H., & Bredenkamp, J. (Eds.), Enzyklopädie der Psychologie: Messen und Testen. Göttingen: Hofgrefe. 257-345.

Schönemann, P. H., Dorcey, T., & Kienapple, K. (1985). Subadditive concatenation in dissimilarity judgments. Perception and Psychophysics, 38, 1-17.

Schönemann, P. H., & Lazarte, A. (1987). Psychophysical maps for subadditive dissimilarity ratings. Perception and Psychophysics, 42, 342-354.

Schroeder, H. M., Driver, M. J., & Streufert, S. (1967). Human information processing. New York, Holt, Rinehart & Winston.

Scott, W. A. (1962). Cognitive complexity and cognitive flexibility. Sociometry, Vol. 25, 405-414.

Scott, W. A. (1963). Cognitive complexity and cognitive balance. Sociometry, 26, 66-74.

Shepard, R. N. (1957). Stimulus and response generalization: a stochastic model relating generalization to distance in psychological space. Psychometrika, Vol. 22, 325-345.

Shepard, R. N. (1958). Stimulus and response generalization: deduction of the generalization gradient from a trace model. Psychological Review, Vol. 65, 242-256.

Shepard, R. N. (1960). Similarity of stimuli and metric properties of behavioral data. In: Gulliksen, H., & Messick, S. (Eds.), Psychological scaling: Theory and applications. New York, Wiley, 33-45.

Shepard, R. N. (1962). The analysis of proximities: multidimensional scaling with an unknown distance function. Psychometrika, Vol. 27, 125-139; 219-246

Shepard, R. N. (1963). Analysis of proximities as a technique for the study of information processing in man. Human Factors, Vol. 5, 33-48.

Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. Journal of Mathematical Psychology, Vol. 1, 54-87.

Shepard, R. N. (1965). Multidimensional scaling of similarity. Psychometrika, Vol. 30, 379-393.

Shepard, R. N. (1966). Metric structures in ordinal data. Journal of Mathematical Psychology, Vol. 3, 287-315.

Shepard, R. N. (1972). Psychological representation of speech sounds. In David, E., & Denes, P. (Eds.), Human communication: A unified view. New York: McGraw-Hill.

Shepard, R. N. (1974). Representation of structure in similarity data: problems and prospects. Psychometrika, Vol. 39, No. 4, 373-421.

Shikiar, R. (1974). The perception of politicians and political issues: A multidimensional scaling approach. Multivariate Behavioral Research, Vol. 9, 461-478.

Silver, C. A., Landis, D., & Messick, S. (1966). Multidimensional analysis of visual form: An analysis of individual differences. American Journal of Psychology, Vol. 79, 62-72.

Society of automotive engineers of Japan, Inc. Noise and vibration committee. (1994). Mimic sound expression for vehicle noise.

Spearman, C. (1927). The abilities of man. New York: Macmillian.

Spence, I. (1972). A Monte Carlo evaluation of three nonmetric multidimensional scaling algorithms. Psychometrika, Vol. 37, 461-486.

Spence, I. (1983). Monte Carlo simulation studies. Applied Psychological Measurement, Vol. 7, No. 4, 405-425.

Spence, I., & Graef, J. (1974). The determination of the underlying dimensionality of an empirically obtained matrix of proximities. Multivariate Behavioral Research, 9, 331-341.

Spence, I., & Ogilvie, J. C. (1973). A table of expected Stress values for random rankings in nonmetric multidimensional scaling. Multivariate Behavioral Research, Vol. 8, No. 3, 511-517.

Staufenbiel, T., & Borg, I. (1987). Dimensionen der Ähnlichkeit einfacher Reize: Breite-Höhe oder Grösse-Form? Psychologische Beiträge, 29, 403-422.

Steinheiser, F. H., Jr. (1970). Individual preference scales within a multidimensional similarities space. Journal of Experimental Psychology, Vol. 86, 325-327.

Stenson, H. H., & Knoll, R. L. (1969). Goodness of fit for random rankings in Kruskal's nonmetric scaling procedure. Psychological Bulletin, Vol. 71, No. 2, 122-126.

Stevens, S. S. (1934). The attributes of tone. Proceedings of the National Academy of Science 20, 457-459.

Stevens, S. S. (1957). On the psychophysical law. Psychological Review, 64, 153-181.

Stewart, T. R. (1974). Generality of multidimensional representations. Multivariate Behavioral Research, Vol. 9, 507-519.

Stumpf, C. Tonpsychologie Leipizg: Hirzel. I (1883); II (1890).

Stumpf, C. (1910). Konsonanz und Konkordanz. Nebst Bemerkungen über Wohlklang und Wohlgefälligkeit musikalischer Zusammenklänge. Zeits. f. Psych., Vol. 58, 321-355.

Susini, P., McAdams, S., & Winsberg, S. (1999). A multidimensional technique for sound quality assessment. Acustica, Vol. 85, 650-656.

Terhardt, E. (1968a). Über ein Äquivalenzgesetz für Intervalle akustischer Empfindungsgrößen. Kybernetik, 5, 127-133.

Terhardt, E. (1968b). Über akustische Rauhigkeit und Schwankungsstärke. Acustica, Vol. 20, 215-224.

Terhardt, E. (1974). On the perception of periodic sound fluctuations (roughness). Acustica, Vol. 30, 201-213.

Terhardt, E. (1979). Calculating virtual pitch. Hearing Research, Vol. 1, 155.

Terhardt, E. (1981). Aspekte und Möglichkeiten der gehörbezogenen Schallanalyse und Bewertung. In Fortschritte der Akustik, DAGA ‘81, 99-110.

Terhardt, E. (1984). Wohlklang und Lärm aus psychophysikalischer Sicht. In: Schick, A. & Walcher, K.-P. (Hrsg.): Beiträge zur Bedeutungslehre des Schalls. Bern: Lang.

Terhardt, E., & Stoll, G. (1978). Bewertung des Wohlklangs verschiedener Schalle. Fortschritte der Akustik. VDE-Verlag, Berlin, 583-587.

Terhardt, E., & Stoll, G. (1981). Skalierung des Wohlklangs (der sensorischen Konsonanz) von 17 Umweltschallen und Untersuchung der beteiligten Hörparameter. Acustica, Vol. 48, 215-253.

Terhardt, E., Stoll, G., & Seewann, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. J. Acoust. Soc. Am., 71(3), 679-688.

Thurstone, L. L. (1927). Law of comparative judgment. Psychol. Rev., Vol. 34, 273-286.

Torgerson, W. S. (1952). Multdimensional scaling. I. Theory and method. Psychometrika, Vol. 17, 401-419.

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

Torgerson, W. S. (1965). Multdimensional scaling of similarity. Psychometrika, Vol. 30, 379-393.

Tucker, L. R. (1972). Relations between the multidimensional scaling and three-mode factor analysis. Psychometrika, Vol. 37, 3-27.

Tucker, L. R. & Messick, S. (1963). An individual differences model for multidmensional scaling. Psychometrika, Vol. 28, 333-367.

Tversky, A. (1977). Features of similarity. Psychological Review, 84, 327-352.

Tversky, A., & Gati, I. (1982). Similarity, separability, and the triangle inequality. Psychological Review, 89, 123-154.

Tversky, A., & Krantz, D. H. (1969). Similarity of schematic faces: A test of interdimensional additivity. Perception & Psychophysics, Vol. 5, 124-128.

Tversky, A., & Krantz, D. H. (1970). The dimensional representation and the metric structure of similarity data. Journal of Mathematical Psychology, Vol. 7, 572-596.

Tversky, A., & Krantz, D. H. (1975). Similarity of rectangles: An analysis of subjective dimensions. Journal of Mathematical Psychology, Vol. 12, 4-34.

Van der Auweraer, H., Wyckaert, K., & Hendrix, W. (1997). From sound quality to the engineering of solutions for NVH problems: Case studies. Acustica, Vol. 83, 796-804.

Völk, W. (1981). Lautheit und Lästigkeit von Geräuschen. In Schick, A. (Ed.). Akustik zwischen Physik und Psychologie, Klett-Cotta, Stuttgart, 63-67.

Vogel, A. (1974). Ein gemeinsames Funktionsschema zur Berechnung der Lautheit und der Rauhigkeit. Dissertation. Technical University of München.

Vogel, A. (1975). Über den Zusammenhang zwischen Rauhigkeit und Modulationsgrad. Acustica, Vol. 32, 300-306.

Wagenaar, W. A., & Padmos, P. (1971). Quantitative interpretation of Stress in Kruskal's multidimensional scaling technique. British Journal of Mathematical and Statistical Psychology, Vol. 24, 101-110.

Wainer, H., & Kaye, K. (1974). Multidimensional scaling of concept learning in an introductory course. Journal of Educational Psychology, Vol. 66, 591-598.

Walters, H. A., & Jackson, D. N. (1967). Group and individual regularities in trait inference: a multidimensional scaling analysis. Multiv. Behav. Res., Vol. 2, 537-541.

Weber, E. G., & Zener, K. E. (1928). The method of absolute judgment in psychophysics. Psychological Review, 35, 466-493.

Weber, R. (1990). Psychoacoustical parameters contributing to the unpleasantness of noise. In Müller, F. (Hrsg.). Fechner Day 90. Proceedings of the Sixth International Society for Psychophysics, Würzburg, 193-198.

Wellek, A. (1975). Musikpsychologie und Musikaesthetik. Grundriss der Systematischen Musikwissenschaft. Bouvier Verlag Herbert Grundmann. Bonn.

Wender, K. (1969). Die psychologische Interpretation nichteuklidischer Metriken in der multidimensionalen Skalierung. Unpublished doctoral dissertation. University of Darmstadt.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. Psychologische Forschung, 4, 301-350.

Widmann, U. (1992). Ein Modell der Psychoakustischen Lästigkeit von Schallen und seine Anwendung in der Praxis der Lärmbeurteilung. Unpublished doctoral dissertation. Technical University of München.

Widmann, U. (1997). Three application examples for sound quality design using psychoacoustic tools. Acustica, Vol. 83, 819-826.

Widmann, U., & Fastl, H. (1998). Calculating roughness using time-varying specific loudness spectra. In: Davies, P., & Ebbitt, G. (Eds.). Sound Quality Symposium, Ypsilanti Michigan USA, 55-60.

Winsberg, S., & Carroll, J. D. (1989). A quasi-nonmetric method for multidimensional scaling via an extended Euclidean model. Psychometrika, Vol. 54, 217-229.

Winsberg, S., & De Soete, G. (1993). A latent class approach to fitting the weighted Euclidean model, CLASCAL. Psychometrika, Vol. 58, 315-330.

Wish, M., Deutsch, M., & Biener, L. (1970). Differences in conceptual structures of nations: An exploratory study. Journal of Personality and Social Psychology, 16, 361-373.

Wolfrum, C. (1976). Zum Auftreten quasiäquivalenter Lösungen bei einer Verallgemeinerung des Skalierungsverfahrens von Kruskal auf metrische Räume mit einer Minkowski-Metrik. Archiv für Psychologie, 128, 96-111.

Young, F. W. (1970). Nonmetric multidimensional scaling: recovery of metric information. Psychometrika, Vol. 35, 455-473.

Young, F. W., & Torgerson, W. S. (1967). TORSCA, a Fortran IV program for Shepard-Kruskal multidimensional scaling analysis. Behavioral Science, Vol. 12, 498.

Young, G., & Householder, A. S. (1938). Discussion of a set of points in terms of their mutual distances. Psychometrika, Vol. 3, 19-22.

Yu, W. (1987). Accuracy of measurement for assessing human response on noise. In: Proceedings 1987 International Conference on Noise Control Engineering. Inter'Noise 87 Beijing. Acoustical Society of China, Vol. II, 985-988.

Zajonc, R. B. (1980). Feeling and thinking. Preferences need no inferences. American Psychologist, Vol. 35, No. 2, 151-175.

Zwicker, E. (1958). Über psychologische und methodische Grundlagen der Lautheit. Acustica, Vol. 8, 237-258.

Zwicker, E. (1960). Verfahren zur Berechnung der Lautstärke. Acustica, Vol. 10, 304-310.

Zwicker, E. (1982). Psychoakustik. Springer, Berlin.

Zwicker, E. (1991a). Evaluation of sounds (loudness and annoyance). A proposal for defining and calculating the unbiased annoyance. In: Schick, A., Hellbrück, J. & Weber, R. (Eds.): Contributions to Psychological Acoustics V. Results of the 5th Oldenburg Symposium on Psychological Acoustics. Oldenburg: BIS, 187-202.

Zwicker, E. (1991b). Ein Vorschlag zur Definition und zur Berechnung der unbeeinflußten Lästigkeit. Zeitschrift für Lärmbekämpfung, Vol. 38, 91-97.

Zwicker, E., Deuter, K., & Preisl, W. (1985). Loudness meters based on ISO 532 B with large dynamic range. Proc. Internoise ‘85, Vol. II, 1119-1122.

Zwicker, E., & Fastl, H. (1999). Psychoacoustics. Springer-Verlag, Berlin Heidelberg.

Zwicker, E. & Terhardt, E. (1980). Analytical expressions for critical band rate and critical band width as a function of frequency. J. Acoust. Soc. Amer., Vol. 68, 1523-1525.

Appendix

Erklärung

Ich versichere, daß ich diese Arbeit selbstständig verfaßt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Oldenburg, den 10.02.2001 Byongho Choe