ADA058423

AD No. ______
DDC FILE COPY

(12)

LEVEL II

FEATURE SELECTION IN AUDITORY PERCEPTION

James H. Howard, Jr. and James A. Ballas

ONR CONTRACT NUMBER N00014-75-C-0308

Technical Report ONR-78-5

Human Performance Laboratory
Department of Psychology
The Catholic University of America

July, 1978

78 08 31 046

(14) TR-78-5-ONR

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ONR-78-5
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): (6) FEATURE SELECTION IN AUDITORY PERCEPTION,
5. TYPE OF REPORT & PERIOD COVERED: (9) Technical Rep[redacted]t,
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): (10) James H. Howard, Jr. [redacted] James A. Ballas
8. CONTRACT OR GRANT NUMBER(s): (15) N00014-75-C-0308
9. PERFORMING ORGANIZATION NAME AND ADDRESS: The Catholic University of America, Washington, D. C. 20064
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 197-027
11. CONTROLLING OFFICE NAME AND ADDRESS: Engineering Psychology Programs, Code 455, Office of Naval Research
12. REPORT DATE: (11) Jul[redacted]78
13. NUMBER OF PAGES: ~~30~~ (12) 36p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number):

Auditory perception
Feature selection
Feature extraction
Auditory pattern recognition

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Feature extraction plays a fundamental role in most theories of pattern recognition, but despite its importance, the extraction process is not well defined. Two contrasting views of feature extraction can be identified, one which emphasizes invariant feature detection and one which emphasizes flexible feature selection. The invariant detector approach assumes that the auditory system is equipped with finely tuned feature detectors that respond to specific stimulus properties.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

409381 78 08 31 046

In this view, stimuli are described in terms of property lists of specific features. In contrast, the more flexible, process-oriented approach assumes that the auditory system is equipped with a set of rules and criteria for feature selection. In this view, the important perceptual features reflect the underlying structure of the stimuli. Research on timbre and pitch perception has supported a flexible, process-oriented approach. The flexibility of this approach offers particular advantages in that it can explain the effects of stimulus and task context on performance. Both types of context influence the perception of complex sounds. Stimulus context affects the structure of the stimulus space and consequently the features that would be extracted by a structure preserving transformation. Task context affects the relative importance of features in making similarity judgments and classification decisions. The two approaches to feature extraction have important implications for the development of auditory pattern recognition theory.

The feature extraction process plays a fundamental role in many theoretical treatments of auditory pattern recognition. At some early stage in the recognition process, the perceptual representation of a stimulus is broken down into a set of elementary properties or characteristics. The central role of this stage can be seen in Figure 1.

Insert Figure 1 here

In this characterization of the pattern recognition process, the preliminary analysis stage produces a relatively unprocessed representation of an incoming stimulus. At this point the representation is thought to contain considerable noise and redundancy. The output of this stage is then transformed by the feature extraction stage into a relatively small set of distinctive features--the basic building blocks of the recognition process. As Anderson, Silverstein, Ritz, and Jones (1977) have noted, "Distinctive features are usually viewed as a system for efficient preprocessing, whereby a noisy stimulus is reduced to its essential characteristics and a decision is made on these" (p. 429).

Quite clearly, feature extraction involves information reduction. Some information in a pattern is retained while other information is discarded. Ideally, the set of distinctive features should uniquely specify the stimulus, preserving or enhancing perceptually-important differences among stimuli, and reducing or eliminating perceptually-unimportant differences.

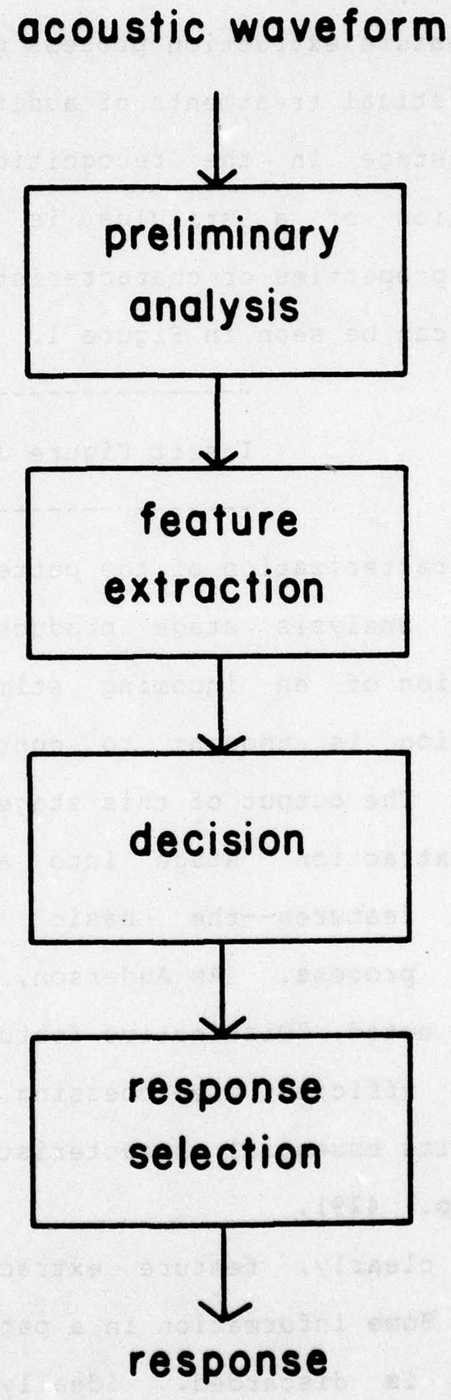

Figure 1. Flow diagram of a four stage pattern recognition model.

The significance of feature extraction in auditory recognition should be obvious. Since the feature representation is efficient both in terms of dimensionality and redundancy, the subsequent decision process can be undertaken with minimal effort and optimal reliability. On the other hand, an ineffective set of distinctive features not only can increase the amount of subsequent processing required, but, by definition, will also make satisfactory performance impossible.

Despite its central importance as a theoretical construct, the feature extraction process has not been well-specified in the literature. No true psychological theory of feature extraction exists. When we say that a stimulus is reduced to its "essential elements," what do we mean? How are these crucial elementary units determined? Implicit in the above discussion is the assumption that a feature tuning process exists whereby a set of distinctive features is defined. In this presentation we focus on possible mechanisms that underlie the feature tuning or feature selection process in human auditory recognition.

<u>Feature Selection Processes</u>

As outlined above, the feature selection problem involves picking a set of distinctive features from the vast set of all possible features. The problem seems clear in the case of a statistician who is constructing an algorithm to classify a set of acoustic patterns. What acoustic cues or combinations of acoustic cues should be considered? Indeed, the ideal features may be some complex function of a number of more primitive spectral measurements. Given a set of preliminary measurements and the desired categorization of the stimuli, the statistician must select a set of distinctive features bearing in mind both performance and economic (i.e., "how much will the feature extraction process cost?") considerations (cf. Meisel, 1972).

The feature selection problem for the psychologist has much in common with that of the statistician. However, instead of actually determining a set of "efficient" distinctive features for human auditory recognition (although this may be of interest in some applied contexts), it is our task to identify the features or the feature selection process that human listeners actually use.

Although a number of specific "natural" auditory feature selection processes can be proposed, two contrasting views are implicit in the literature. The first possibility is that Nature has selected an optimal set of distinctive features through natural selection and built specific mechanisms, finely tuned to detect these features, into our auditory systems. The second view affords more flexibility. Perhaps Nature has built the feature selection process into our auditory systems. In other words, we may have internalized a set of rules and processes that enable us to establish what the distinctive features should be in any particular stimulus context. Both views are considered in more detail below.

The property-list approach. The first view argues that man is equipped with a set of specific feature detecting mechanisms. In terms of auditory pattern recognition, this approach places an emphasis on the feature detectors themselves. An auditory feature extractor is, from this perspective, a filter-like device that monitors the incoming stream of sensory information for particular stimulus properties. In short, each detector is tuned to "look for" a particular stimulus property, and a set of feature detectors determine a property list for the stimulus. In the extreme, this seems consistent with previous neurophysiological investigations of single unit responding in sensory systems. The first relevant evidence emerged many years ago from the seminal work of Lettvin, Maturana, McCulloch, & Pitts (1959) with the frog's visual system. Their pioneering research revealed the presence of highly selective neurons in the visual periphery tuned to select stimuli of particular relevance for the animal's survival. Moving edge, overall illumination, and the highly-popularized "bug" detectors were among those discovered. It seems that Nature equipped the frog with special detectors for virtually everything it needs to know about its visual world. Similar work soon followed, investigating feature detectors in both the visual (Hubel & Wiesel, 1962) and auditory systems (Whitfield & Evans, 1965). This research stimulated considerable speculation about hierarchical decision mechanisms where feature information is combined and re-combined, ultimately leading to classification (<u>cf</u>. Weisstein, 1973). When extended, this line of reasoning leads us to Sherrington's (1941) notion of a supreme, "pontifical" cell whose response signals the presence of a particular complex pattern. Although he opted for a more democratic system of "cardinal" cells in place of the all-knowing pontiff, Barlow (1972) succinctly summarized this approach in his specification of a "neuron doctrine for perceptual psychology."

A good example of this sort of system in human audition is the set of distinctive features and feature detectors hypothesized to underlie human speech perception (Fant, 1973). Here a relatively small number of distinctive features have been described that may be used to uniquely characterize individual phonemes. A voicing detector, for example, would monitor the speech stream for cues that distinguish between voiced and unvoiced stop consonants. In an initial study, Eimas & Corbit (1973) used a psychophysical procedure to obtain evidence for the existence of voicing detectors, finely tuned to a relatively narrow range of voice onset times (i.e., formant onset asynchronies in the speech signal). More recent work has generalized their findings to include the psychophysical investigation of a variety of linguistic as well as non-linguistic feature detectors in the human auditory system (Cooper, 1975).

The process-oriented approach. The second alternative assumes that man has internalized the feature selection process itself. In contrast to the property-list approach, the important "features" in a complex sound reflect whatever structure exists in the output of the preliminary analysis stage. In this sense, Nature has endowed us with a set of rules and criteria for feature selection rather than with highly-tuned detection mechanisms.

In arguing that the feature selection process is built in, we necessarily assume that certain general principles exist that can characterize feature selection across a variety of stimulus and task conditions. These invariants include both the selection criteria employed and a mechanism for applying them. In the present discussion, we assume that the feature selection process attempts to reduce the dimensionality of the stimulus representation while preserving as much of the stimulus structure as possible. An example of this approach may be seen in Wightman's (1973) pattern transformation model of periodicity pitch perception. The model assumes that the auditory system performs two successive Fourier transforms (equivalent to an autocorrelation in the time domain) to extract periodicity information from a complex tone. Since signal periodicity reflects the frequency relations among individual spectral components, the proposed transformation analyzes the relational structure of the stimulus. Uttal (1975) has outlined a similar autocorrelational model for visual pattern detection.

Implications for auditory perception. The intention of this discussion is not to suggest that one or the other of these views is necessarily correct--at this point, the most nearly correct view would seem to include elements of both perspectives. Rather, we consider these particular approaches because they occupy opposite ends on a continuum of feature selection flexibility. This difference has a number of implications for auditory recognition theory, and in particular, for recognition processes in the perception of the timbre of complex sounds. As Plomp (1976) has noted, timbre is generally defined as "that attribute of auditory sensation in terms of which a listener can judge that two steady-state complex tones having the same loudness and pitch, are dissimilar" (Pp. 85-86). In other words, timbre is everything left over after we take away loudness and pitch. Quite clearly, timbre does not describe a single perceptual attribute of sound, but rather, it represents a family of perceptual attributes. One would be entirely at ease in reporting that a complex sound has a high pitch or is very loud, but to attempt a simple description of this sort for timbre would seem ridiculous. If asked to discuss the timbre of a sound, the listener would likely resort to a number of adjectives, "it is coarse, pleasant, bright, etc." (von Bismarck, 1974). But even given this verbal flexibility, the listener would find it difficult to adequately describe the timbre of a complex tone.

If one were to adopt a property-list approach to the feature extraction stage for timbre perception, it would be necessary to specify a list of important timbre attributes. However, it should be clear that any list of possible timbre properties or features would be very long indeed. In this case, the more flexible process-oriented approach to feature extraction would seem most appropriate. Rather than searching for the set of distinctive features that would enable the listener to distinguish all possible timbres, in this approach we attempt to specify a process that can characterize the relations among individual components in the amplitude and phase spectra of complex sounds.

In the remainder of this paper we examine three experiments whose findings suggest that considerable flexibility exists in the feature extraction process. In all three experiments, the features that listeners use in perceiving the timbre of complex sounds were investigated. The results of the first two experiments serve to emphasize the importance of the stimulus population in determining the timbre attributes that listeners use in comparing complex sounds. The findings of the third experiment illustrate the role of task factors in the feature extraction process.

Stimulus Effects in Feature Extraction

In the above discussion we argued that the feature extraction process in timbre perception may be more appropriately viewed as a structure analyzing process than as a feature detection process. In order to evaluate this hypothesis empirically it is necessary to examine the output of the feature extraction stage, and to relate this feature representation to the known properties of its input. Although the feature representation is obviously not directly observable, it may be inferred using a variety of psychophysical procedures. In particular, multidimensional scaling has emerged as a useful method for identifying the underlying psychological structure of complex sounds (Plomp, 1976). Typically, listeners are asked to provide pairwise or triadic dissimilarity judgments on the set of signals of interest. A specific multidimensional scaling algorithm is then applied to decompose the resulting subjective proximity matrix into an n-dimensional metric space in which each stimulus is represented as a single point or vector. Providing that an interpretable solution with satisfactory stress exists, it is generally assumed that the dimensions in the psychological stimulus space reflect those features that the listeners employed in comparing the stimuli. In other words, it is at least implicitly assumed that the scaling methods provide an approximate inverse to the later stages of the recognition process (_cf._ Figure 1). If we are willing to make certain assumptions about the information available after the preliminary analysis stage, then we have the input/output information necessary to speculate about the feature extraction process.

In the first systematic application of these methods to audition, Plomp and his associates (Plomp, 1970) compared the timbre properties of nine musical instrument sounds to their corresponding spectral structure. A three-dimensional stimulus space was revealed in a multidimensional scaling analysis of the subjective similarity data for these stimuli. The configuration of interstimulus distances in this perceptual space correlated highly with the corresponding distances in a three-dimensional physical space obtained in a physical analysis of these sounds. Although Plomp was primarily interested in determining whether a correlation existed, for our present purposes, the specific methods used to obtain the physical space are of particular interest.

Specifically, the physical analysis was based on information that could be reasonably thought to approximate that available to the auditory feature extraction process. Recognizing the limited spectral analyzing ability of the human auditory system (Zwicker, Flottorp, & Stevens, 1957), Plomp obtained 1/3-octave band-level measurements for each of his complex sounds. A principal-components analysis was then performed on these spectra. The results of this analysis revealed that each of the nine stimuli could be characterized in terms of three spectral attributes with very little loss of information. If we are willing to assume that the 1/3-octave spectrum approximates the output of the preliminary analysis stage depicted in Figure 1, then these findings suggest that the listener's feature extraction process may be somewhat similar to a principal-components analysis.

This conclusion is entirely consistent with our hypothesis that the feature extraction stage in timbre perception involves a structural analysis of the stimuli. More specifically, the principal-components analysis may be thought of as a structure preserving transformation that maps stimuli from one space into another of lower dimensionality. The first principal component is simply a new axis in the original space (in this case the measurement space spanned by the 1/3-octave band-levels) that accounts for most of the variability in the data. In other words, the set of projections of stimuli in the measurement space onto the first principal component has maximum variance. The second principal component is an axis orthogonal to the first that accounts for most of the residual variance and so on. [In practice, the principal components are determined by selecting the eigenvectors of the covariance matrix for the stimulus set that correspond to the largest eigenvalues (Harris, 1975)].

In this view, then, the feature extraction process selects a subspace of the original space that preserves as much of the variability among stimuli as possible. It is clear that these features (i.e., the principal components) reflect the structure of the stimulus set, and therefore we would expect the important perceptual features to vary dramatically depending on the stimulus context.

This finding is given some generality by a similar result obtained in our laboratory. In our experiment listeners were asked to rate the pairwise similarity of eight passive sonar recordings. Two perceptual dimensions were extracted from these data using a metric multidimensional scaling procedure (Howard, 1977). The results were then compared with the outcome of a physical analysis of the stimuli that paralleled the analysis described above. The 1/3-octave spectrum was obtained for each of the eight sounds. These data were then submitted to a principal-components analysis. Since most of the variability could be accounted for by the first principal component, it was concluded that the steady-state characteristics of these sounds could be adequately summarized by a single measurement. This derived physical attribute closely approximated one of the perceptual dimensions obtained in the scaling analysis. (The other perceptual dimension revealed in the scaling analysis reflected a temporal property of the sounds and is not directly relevant to this discussion.) A closer examination of the specific signal values on this extracted dimension suggested that it summarized the overall shape of the spectra, and in particular, the degree of bimodality of the spectra. When asked to describe stimulus differences along this dimension, listeners used such terms as "this one is more uniform" or "in this one there seems to be more than one sound present."

In this experiment, as in Plomp's, it appears that a structure preserving transformation reasonably approximates the analysis performed by the feature extraction stage.[1] Similar findings have also been reported for the analysis of steady-state vowel spectra (Klein, Plomp, & Pols, 1970). Although these experiments were conducted for another purpose, the findings are generally consistent with the present hypothesis that the feature extraction stage for timbre perception is best viewed as a structure analyzing process. Since the principal components are simply weighted linear combinations of the more basic measurements (in this case the 1/3-octave band-levels), it is clear that we could also develop a weighted property-list scheme to account for these findings. In such a system, the listener would adjust the measurement weights to develop "features" that maximally discriminate among the stimuli. Nonetheless, the objectives of this system are more naturally discussed in terms of the structure analyzing approach.

The major emphasis of the above discussion is that feature extraction is both efficient, in a dimensionality reducing sense, and flexible, in that it readily adapts to the stimulus context. We have argued that a principal-components analysis shares these characteristics and is therefore a possible model for the feature extraction process in timbre perception. We would like to emphasize, however, that a variety of other structure preserving transformations are also adequate. For example, we could select a multidimensional scaling algorithm that would reduce the dimensionality while preserving the configuration of inter-stimulus distances in the measurement space.

__Task Effects in Feature Extraction__

In the experiments outlined above, listeners were simply required to evaluate relative stimulus similarity. In this situation there are no correct or incorrect judgments. It therefore seems reasonable that listeners would employ a feature extraction transformation that preserves as much of the spectral information in the stimuli as possible. It is obvious, however, that in a classification task where performance is evaluated in terms of external criteria, the requirements of the feature extraction process would be quite different. As Figueiredo (1976) has pointed out, the performance of the entire system must be considered when selecting the optimum features in this situation. We have already seen in Getty's paper (Getty, Swets, Swets, & Green, in press) that observers emphasized different features in a visual classification task than they did in a visual comparison task. The experiment described below (Howard, Ballas, & Burgy, 1978) demonstrates a similar effect in an aural classification situation, and illustrates the role of task factors in feature extraction.

The stimuli investigated in this study consisted of sixteen broadband noise signals amplitude modulated by sawtooth waves of varying frequency and attack. Four levels of modulation frequency (4, 5, 6, and 7 Hz) and four combinations of attack/decay (20 and 40 msec) were used. Eight listeners learned to classify the sixteen signals on the basis of one of two eight-category partitions. The two partitions were selected to emphasize one or the other dimension by requiring listeners to discriminate among all four levels of this dimension and only two levels of the other dimension. The two partitions are presented schematically in Figure 2. Here we have labeled the perceptual dimension corresponding to attack "Quality, " and the perceptual dimension corresponding to modulation frequency "Tempo."

Insert Figure 2 here

Clearly, the features are not of equal importance in the two partitions. Listeners in the Quality group were required to discriminate relatively small differences in attack, whereas listeners in the Tempo group were required to discriminate relatively small differences in modulation frequency. The confusion data from this experiment were analyzed in terms of a probabilistic model of the classification process.

Our model assumes that the decision stage operates on the output of the feature extraction process and classifies stimuli so as to maximize the probability correct (cf. Howard et al, 1978). An important assumption of the model is that, with feedback, listeners perform a selective tuning of their feature extraction processes. Theoretically, the tuning process determines a weighting factor for each of the two features.[2] Selective tuning occurs when the weighting factor for one

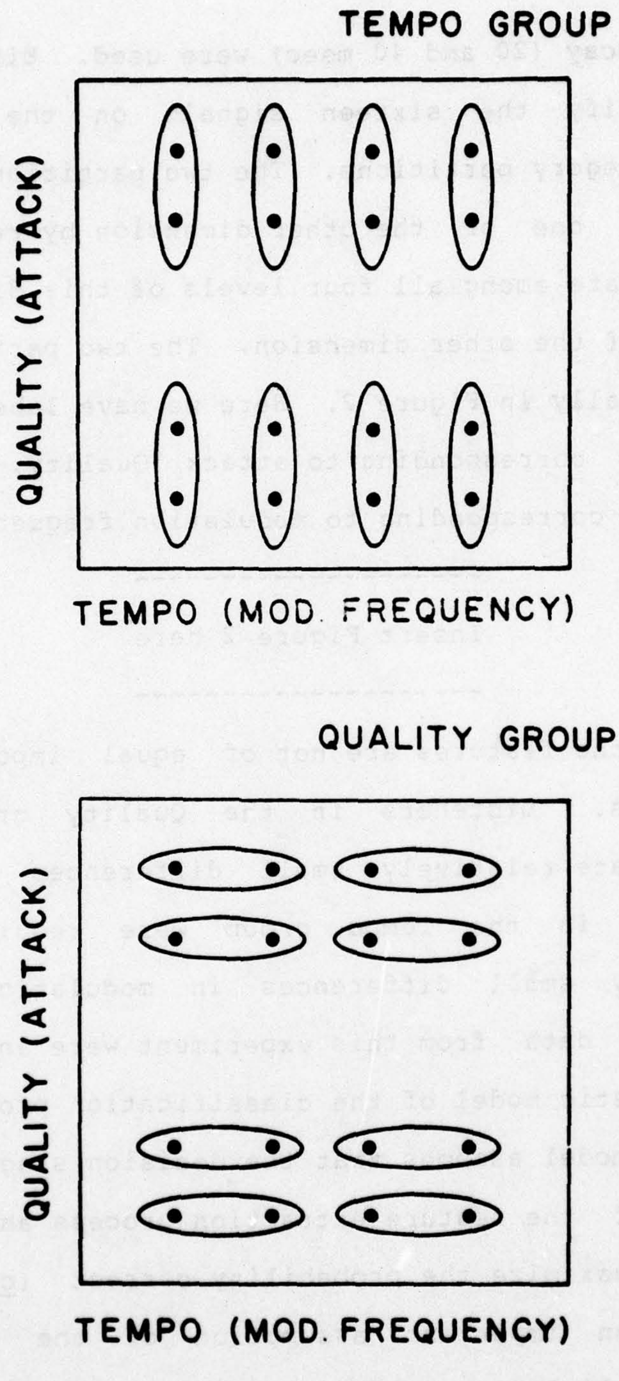

Figure 2. Schematic representation of two partitions of the sixteen signals into eight categories.

dimension increases relative to the other.

Weighting parameters for each feature were estimated for individual listeners by fitting the model to the observed confusion matrices using a standard gradient technique. The weights obtained for each practiced listener are displayed in Figure 3.

Insert Figure 3 here

It is evident that our listeners responded to the demands of their classification task. Listeners in the Tempo group had greater weights for signal Tempo than signal Quality, whereas the opposite was true for the Quality group. It may also be noted that no individual, with the possible exception of listener PH, maximized the weights for both features simultaneously. We interpreted this finding to suggest that, at least in the context of this experiment, the total amount of feature tuning that can occur at any point is limited. Since we assumed that feature tuning reflects the operation of a selective attentional mechanism, this interpretation is consistent with recent limited capacity views of human attention (e.g., Kahneman, 1973).

Given an overall limit on the amount of feature tuning that can occur, we wondered what strategy the listeners used to determine how much emphasis to place on each feature. In our decision model we assumed that listeners attempt to maximize the probability correct. Therefore, we investigated the possibility that the feature weights were also determined on this basis by

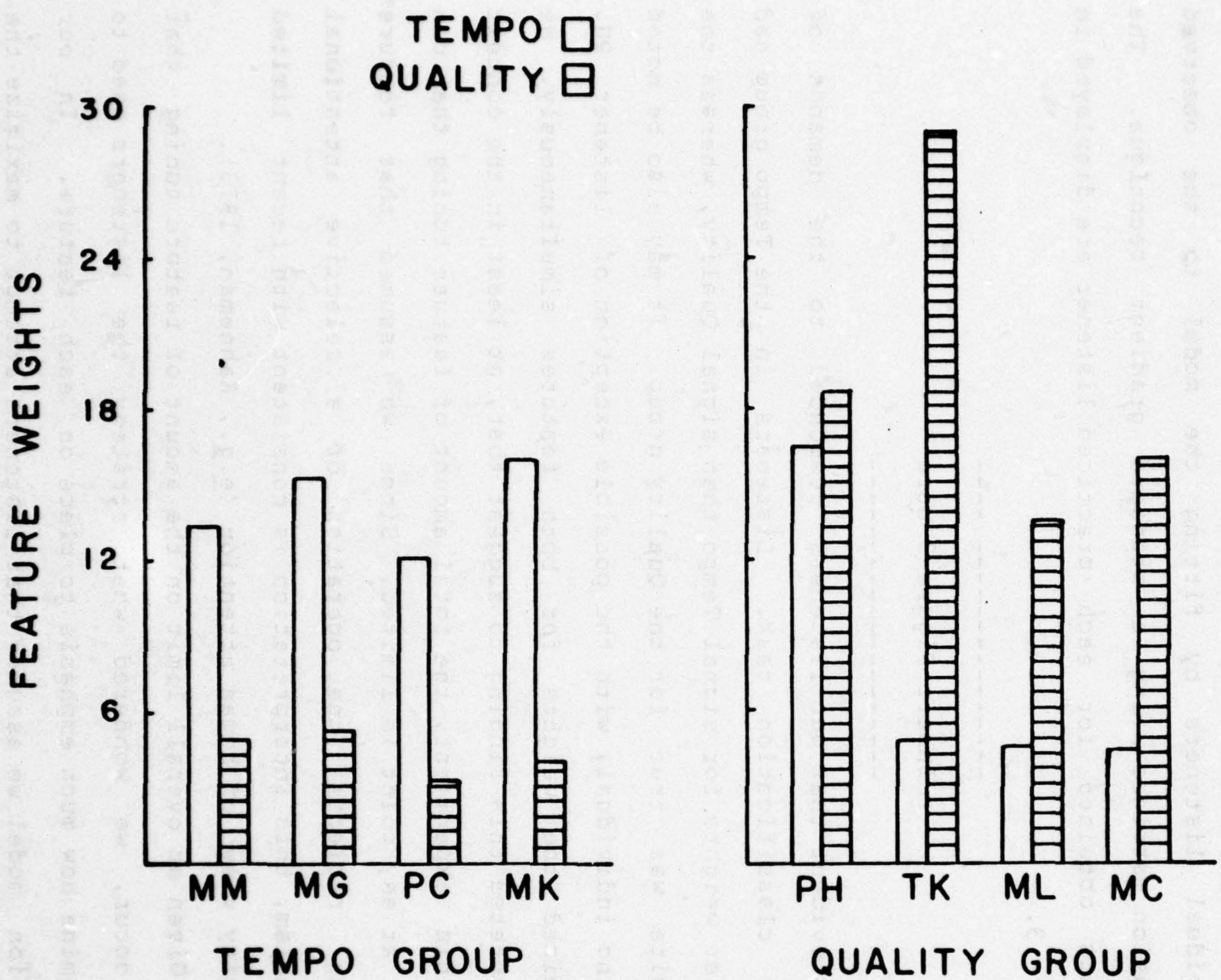

Figure 3. Estimated weighting parameters for each feature by individual listener and group.

estimating the theoretically optimal weights for each practiced listener. In computing these values, the sum of the two observed weights was taken to reflect the overall attentional effort expended by each listener. This overall value was then apportioned between the two features so as to maximize the average probability correct. The normalized obtained and optimal weights reflecting the relative importance of signal Tempo for each listener are displayed in Figure 4.

Insert Figure 4 here

Although in general the obtained weights are reasonably well approximated by the optimal values (the overall Pearson product-moment correlation, $\underline{r}(15) = .98$), a small but consistent discrepancy is evident. Six of our eight listeners showed a tendency to overemphasize the more important of the two features.

Nonetheless, it is clear from these findings that listeners show considerable flexibility in the emphasis they place on individual perceptual features in an aural classification task. With practice, feature tuning processes increase the importance of both features. More importantly, for experienced listeners the tuning process appears to operate selectively with relative feature emphasis determined by a strategy that attempts to maximize the overall probability correct. Of further interest is a comparison of these findings with the results of two multidimensional scaling studies involving the stimuli described above. In the first, an independent group of 30 listeners rated

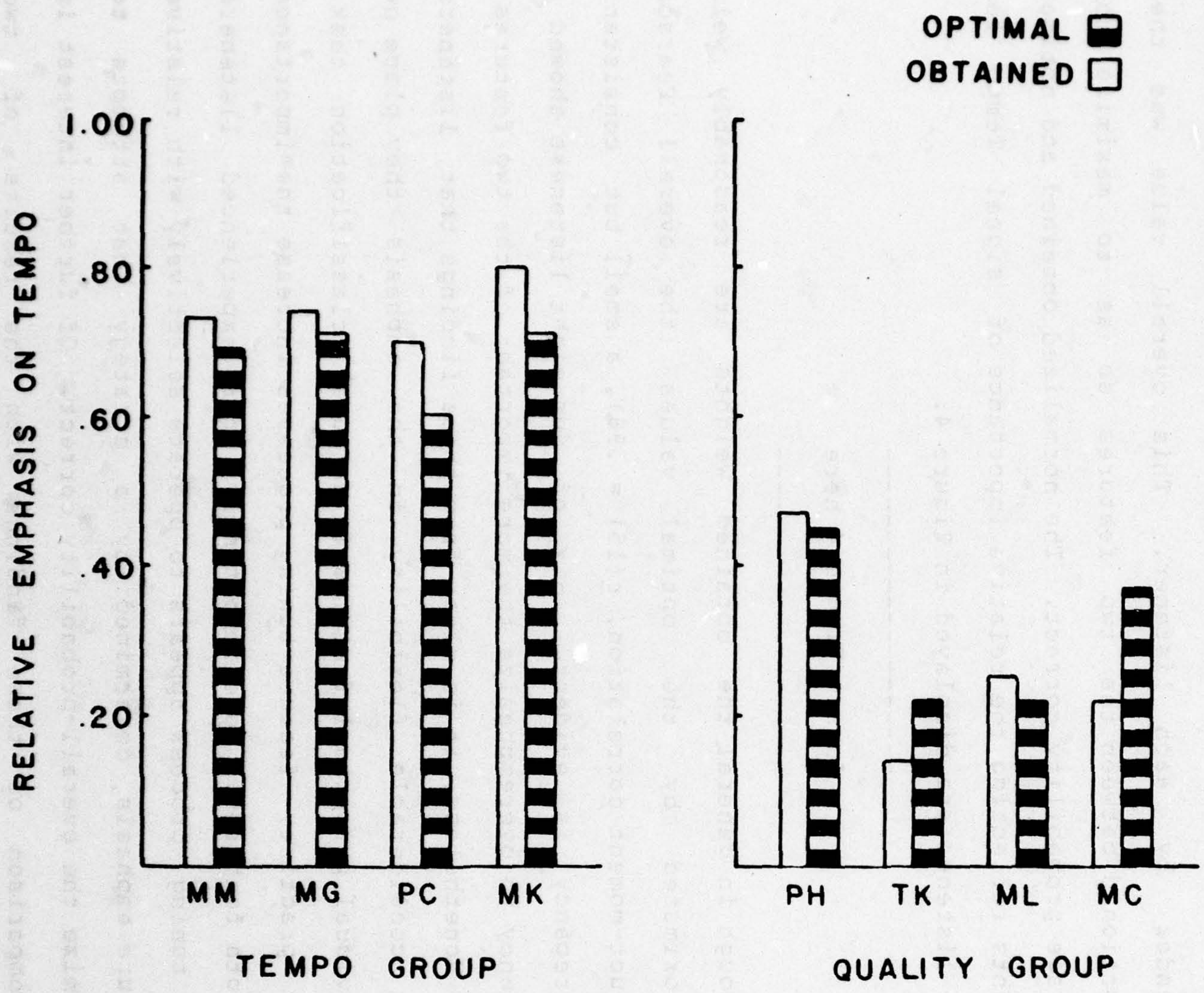

Figure 4. Normalized obtained and optimal weights for the Tempo feature by individual listener and group.

the pairwise similarity of all sixteen signals. A multidimensional scaling analysis of these data revealed that 22 of the 30 participants placed a greater emphasis on signal Quality than Tempo. In the second experiment, the observers from the above classification experiment rated the pairwise similarity of the eight stimulus categories they had learned. In this case, only the category labels were provided, no signals were actually presented. The data were decomposed into a "conceptual" space where the dimensions corresponded to modulation frequency (_i.e._, Tempo) and attack (_i.e._, Quality). Although we had expected the subjective importance of these dimensions to be strongly influenced by the prior classification training (_i.e._, Tempo more important for the Tempo group and Quality more important for the Quality group), listeners in both groups placed a greater relative emphasis on signal Quality. This finding is more in line with the results of the first scaling study than with the results of the classification study. It appears, therefore, that in the similarity rating task, listeners tended to emphasize Quality more than Tempo, whereas in the classification task they emphasized the dimension that would lead to optimal performance. This result further illustrates the importance of task factors in determining the relative subjective importance of stimulus features.

Conclusion

As indicated in the introduction, the feature extraction process plays a central role in most theoretical approaches to auditory pattern recognition. Incoming stimuli are analyzed in terms of a set of distinctive or characteristic features that form the basis for all subsequent perceptual processing. Nevertheless, relatively little research has focused on the selection and tuning processes whereby these essential perceptual properties of stimuli are defined. In this paper we have considered two contrasting views of the feature selection process. On the one hand, it is possible that finely-tuned stimulus property analyzers exist in the human auditory system. As Fant (1973) has pointed out, this property-list or detector-oriented approach is by definition context free. Some property analyzers will respond to a specific stimulus whereas others will not. In contrast, it is also possible that the feature extraction process is highly context-sensitive. In this latter process-oriented approach, feature selection is viewed as a continuous on-going process. The distinctive features used to characterize a particular sound emerge from a structural analysis of the more basic psychophysical measurements obtained by the auditory system.

Overall, the findings outlined above appear consistent with the more flexible, process-oriented approach to feature selection. In the comparative judgment task listeners are required to evaluate stimulus similarity. Since no specific comparison criteria are typically indicated, listeners need only know something of the structure of the stimulus set in order to make their judgments. The features identified in the similarity rating experiments outlined above generally reflect the spectral characteristics of the stimuli. We argued that in these studies, the feature extraction process is most naturally viewed as an on-going structural analysis of the low-resolution spectra extracted from the stimuli by peripheral auditory mechanisms. However, in the case of a classification task, a simple structural analysis of the stimulus set is not sufficient. In this situation, the assignment of stimuli to categories effectively changes the important structural properties of the stimulus set. Particular partitions of the stimuli serve to emphasize some stimulus relations while de-emphasizing others. We have argued that in this sort of task an additional feature tuning process occurs that adjusts the relative emphasis placed on the important structural features to accommodate the external task constraints. Furthermore, the results of our classification study suggested that the feature tuning process operates on a limited-capacity basis, and that the fine-grained adjustment of feature emphasis is based on a strategy that attempts to optimize the probability correct. These findings emphasize the importance of overall performance considerations in the feature extraction process. In this sense, distinctive auditory features are tuned not only to the stimuli, but also to the decision rule employed by the listener (Figueiredo, 1976).

Before closing we would like to offer a few caveats. First, although our conclusions were derived from the findings summarized above, these experiments were not designed to explicitly test the issues addressed in this paper. For this reason our conclusions must be regarded as tentative and speculative. No single experiment, for example, has enabled us to simultaneously examine the effects of both stimulus structure and task demands. Similarly, we have not addressed the detailed problem of how the feature selection process may operate on a trial by trial basis. How might the proposed structural analysis proceed in an incremental fashion? We have argued that in many cases auditory distinctive features are determined largely by the stimulus context. What external and subjective factors determine the relevant context in any particular situation? Quite clearly the answers to these and other questions must await further investigation. Experiments designed to address some of these issues are presently underway in our laboratory.

Finally, it is important to remember that we are not proposing that either of the two approaches we have considered is necessarily "correct." As we have indicated, these approaches represent extremes on a continuum of possible feature selection mechanisms. Although one may challenge the extreme property-list position on logical grounds (e.g., Weisstein, 1973; Uttal, 1978), a weighted property-list approach begins to resemble the process-oriented view discussed above. Furthermore, we suspect that it is impossible to distinguish between a modified property-list model and the process-oriented model using only psychophysical techniques. The distinction we have considered is significant because of its impact on theory and, hence, on the empirical questions that are appropriate to ask. A strict property-list model would direct us to search for evidence regarding invariant auditory feature detectors, whereas a process-oriented model would have us look for common principles underlying feature extraction across a wide range of different stimuli and tasks. Regardless of the specific view that ultimately emerges as a primary solution to the feature selection problem, it is clear that future psychological research in auditory pattern recognition must address the fundamental question of how distinctive auditory features are determined.

Acknowledgements

This work was supported by a contract from the Engineering Psychology Programs, Office of Naval Research. James H. Howard, Jr. is the principal investigator. The authors gratefully acknowledge Donald C. Burgy's contribution to this work, and the helpful comments of Darlene V. Howard, Bruce M. Ross, and James E. Youniss on an earlier version of this manuscript.

References

Anderson, J.A., Silverstein, J.W., Ritz, S.A., & Jones, R.S. Distinctive features, categorical perception, and probability learning: Some applications of a neural model. Psychological Review, 1977, 84, 413-451.

Barlow, H.B. Single units and sensation: A neuron doctrine for perceptual psychology? Perception, 1972, 1, 371-394.

Cooper, W.E. Selective adaptation to speech. In F. Restle, R.M. Shiffrin, N.J. Castellan, H.R. Lindman, & D.B. Pisoni (Eds.), Cognitive theory, (vol. I), Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.

Eimas, P.D. & Corbit, J.D. Selective adaptation of linguistic feature detectors. Cognitive Psychology, 1973, 4, 99-109.

Fant, G. Speech sounds and features. Cambridge: MIT Press, 1973.

Figueiredo, R.J.P. Design of optimal feature extractors by mathematical programming techniques. In C.H. Chen (Ed.) Pattern recognition and artificial intelligence. New York: Academic Press, 1976.

Getty, D.J., Swets, J.A., Swets, J.B., & Green, D.M. On the prediction of confusion matrices from similarity judgments. (Technical Report No. 3719). Cambridge, MA: Bolt, Beranek, and Newman, in press.

Harris, R.J. A primer of multivariate statistics. New York: Academic Press, 1975.

Howard, J.H., Jr. Psychophysical structure of eight complex underwater sounds. Journal of the Acoustical Society of

America, 1977, 62, 149-156.

Howard, J.H., Jr., Ballas, J.A., & Burgy, D.C. Feature selection in auditory perception. (Technical Report ONR-78-4). Washington, D.C.: The Catholic University of America, Human Performance Laboratory, July, 1978.

Hubel, D.H. & Wiesel, T.N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. Journal of Physiology, 1962, 160, 106-154.

Kahneman, D. Attention and effort. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Klein, W., Plomp, R., & Pols, L.C.W. Vowel spectra, vowel spaces, and vowel identification. Journal of the Acoustical Society of America, 1970, 48, 999-1009.

Lettvin, J.Y., Maturana, H.R., McCulloch, W.S., & Pitts, W.H. What the frog's eye tells the frog's brain. Institute of Radio Engineers Proceedings, 1959, 47, 1940-1951.

Meisel, W.S. Computer-oriented approaches to pattern recognition. New York: Academic Press, 1972.

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

Plomp, R. Aspects of tone sensation. New York: Academic Press, 1976.

Sherrington, C.S. Man on his nature. Cambridge: Cambridge University Press, 1941.

Uttal, W.R. An autocorrelation theory of form detection.

Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.

Uttal, W.R. The psychobiology of mind, Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.

von Bismarck, G. Timbre of steady sounds: A factorial investigation of its verbal attributes. Acoustica, 1974, 30, 146-159.

Weisstein, N. Beyond the yellow-volkswagen detector and the grandmother cell: A general strategy for the exploration of operations in human pattern recognition. In R.L. Solso (Ed.) Contemporary issues in cognitive psychology: The Loyola symposium. Washington: V.H. Winston & Sons, 1973.

Whitfield, I.C. & Evans, E.F. Responses of auditory cortical neurons to stimuli of changing frequency. Journal of Neurophysiology, 1965, 28, 655-672.

Wightman, F.L. The pattern-transformation model of pitch. Journal of the Acoustical Society of America, 1973, 54, 407-416.

Zwicker, E., Flottorp, G., & Stevens, S.S. Critical band width in loudness summation, Journal of the Acoustical Society of America, 1957, 29, 548-557.

Footnotes

[1]It should be emphasized that Plomp and his associates (personal communication, 1978) did not refer to the dimensions extracted in their perceptual analysis as "features." Rather, as indicated earlier, they were primarily interested in determining the degree of correlation between the configuration of stimuli in the perceptual space and the steady-state spectra of the sounds.

[2]In our original presentation we actually represent the tuning process as an adjustment in the variability of stimulus likelihood functions. The weighting factors discussed here are inversely related to the estimated standard deviations along each dimension.

DIRECTOR, ENGINEERING PSYCHOLOGY
PROGRAMS, CODE 455
OFFICE OF NAVAL RESEARCH
800 NORTH QUINCY STREET
ARLINGTON, VA 22217

DEFENSE DOCUMENTATION CENTER
CAMERON STATION
ALEXANDRIA, VA 22314

DR. STEPHEN J. ANDRIOLE
ACTING DIRECTOR
CYBERNETICS TECHNOLOGY OFFICE
ADVANCED RESEARCH PROJECTS AGENCY
1400 WILSON BLVD
ARLINGTON, VA 22209

CDR P. A. CHATELIER
OAD(E&LS) OUSDR&E
PENTAGON, ROOM 3D129
WASHINGTON, D.C. 20301

DIRECTOR, ACOUSTIC TECHNOLOGY
PROGRAM, CODE 222
OFFICE OF NAVAL RESEARCH
800 NORTH QUINCY STREET
ARLINGTON, VA 22217

DIRECTOR, PHYSIOLOGY PROGRAM
CODE 441
OFFICE OF NAVAL RESEARCH
800 NORTH QUINCY STREET
ARLINGTON, VA 22217

PROGRAM DIRECTOR
OCEAN TECHNOLOGY
NAVAL OCEAN RESEARCH & DEVELOPMENT
(NORDA)
BAY ST. LOUIS, MS 39529

COMMANDING OFFICER
ONR BRANCH OFFICE
ATTN: DR. J. LESTER
BUILDING 114, SECTION D
666 SUMMER STREET
BOSTON, MA 02210

COMMANDING OFFICER
ONR BRANCH OFFICE
ATTN: DR. CHARLES DAVIS
536 SOUTH CLARK STREET
CHICAGO, IL 60605

COMMANDING OFFICER
ONR BRANCH OFFICE
ATTN: DR. E. GLOYE
1030 EAST GREEN STREET
PASADENA, CA 91106

DR. BRUCE MC DONALD
OFFICE OF NAVAL RESEARCH
SCIENTIFIC LIAISON GROUP
AMERICAN EMBASSY, ROOM A-407
APO SAN FRANCISCO 96503

DIRECTOR
NAVAL RESEARCH LABORATORY
TECHNICAL INFORMATION DIVISION
CODE 2627
WASHINGTON, D.C. 20375

NAVAL RESEARCH LABORATORY
ATTN: CODE 5707
WASHINGTON, D.C. 20375

OFFICE OF THE CHIEF OF NAVAL
OPERATIONS, OP987H
PERSONNEL LOGISTICS PLANS
DEPARTMENT OF THE NAVY
WASHINGTON, D. C. 20350

DR. ANDREAS B. RECHNITZER
NAVAL POSTGRADUATE SCHOOL
OCEAN ENGINEERING
MONTERY, CA 93940

MR. ARNOLD RUBINSTEIN
NAVAL MATERIAL COMMAND
NAVMAT 08T24
DEPARTMENT OF THE NAVY
WASHINGTON, D.C. 20360

COMMANDER
NAVAL AIR SYSTEMS COMMAND
HUMAN FACTORS PROGRAMS, AIR 340F
WASHINGTON, D.C. 20361

COMMANDER
NAVAL ELECTRONICS SYSTEMS COMMAND
HUMAN FACTORS ENGINEERING BRANCH
CODE 4701
WASHINGTON, D.C. 20360

A. E. BISSON
CODE 1939
NSRDC
CARDEROCK, MD 20084

LCDR T. W. SCHROPP
NAVAL SEA SYSTEMS COMMAND
NAVSEA OOC-DA
WASHINGTON, D. C. 20362

MR. JAMES JENKINS
NAVAL SEA SYSTEMS COMMAND
CODE 06H1-3
WASHINGTON, D.C. 20362

DR. JAMES CURTIN
NAVAL SEA SYSTEMS COMMAND
PERSONNEL & TRAINING ANALYSES
NAVSEA 074C1
WASHINGTON, D.C. 20362

DIRECTOR
BEHAVIORAL SCIENCES DEPARTMENT
NAVAL MEDICAL RESEARCH INSTITUTE
BETHESDA, MD 20014

DR. GEORGE MOELLER
HUMAN FACTORS ENGINEERING BRANCH
SUBMARINE MEDICAL RESEARCH LAB.
NAVAL SUBMARINE BASE
GROTON, CT 06340

MR. PHILLIP ANDREWS
NAVAL SEA SYSTEMS COMMAND
NAVSEA 0341
WASHINGTON, D. C. 20362

BUREAU OF NAVAL PERSONNEL
SPECIAL ASSISTANT FOR RESEARCH
LIAISON
PERS-OR
WASHINGTON, D. C. 20370

NAVY PERSONNEL RESEARCH AND
DEVELOPMENT CENTER
MANAGEMENT SUPPORT DEPARTMENT
CODE 210
SAN DIEGO, CA 92152

DR. FRED MUCKLER
NAVY PERSONNEL RESEARCH AND
DEVELOPMENT CENTER
MANNED SYSTEMS DESIGN, CODE 311
SAN DIEGO, CA 92152

MR. A. V. ANDERSON
NAVY PERSONNEL RESEARCH AND
DEVELOPMENT CENTER
CODE 302
SAN DIEGO, CA 92152

LCDR P. M. CURRAN
HUMAN FACTORS ENGINEERING BRANCH
CREW SYSTEMS DEPT., CODE 4021
NAVAL AIR DEVELOPMENT CENTER
JOHNSVILLE
WARMINSTER, PA 18950

LCDR WILLIAM MORONEY
HUMAN FACTORS ENGINEERING BRANCH
CODE 1226
PACIFIC MISSILE TEST CENTER
POINT MUGU, CA 93042

HUMAN FACTORS SECTION
SYSTEMS ENGINEERING
TEST DIRECTORATE
U.S. NAVAL AIR TEST CENTER
PATUXENT RIVER, MD 20670

DR. JOHN SILVA
MAN-SYSTEM INTERACTION DIVISION
CODE 823
NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CA 92152

HUMAN FACTORS ENGINEERING BRANCH
NAVAL SHIP RESEARCH AND
DEVELOPMENT CENTER
ANNAPOLIS DIVISION
ANNAPOLIS, MD 21402

COMMANDING OFFICER
NAVAL COASTAL SYSTEMS LAB
PANAMA CITY, FL 32401

DR. ROBERT FRENCH
NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CA 92152

DR. JERRY C. LAMB
DISPLAY BRANCH
CODE TD112
NAVAL UNDERWATER SYSTEMS CENTER
NEW LONDON CT 06320

NAVAL TRAINING EQUIPMENT CENTER
ATTN: TECHNICAL LIBRARY
ORLANDO, FL 32813

HUMAN FACTORS DEPARTMENT
CODE N215
NAVAL TRAINING EQUIPMENT CENTER
ORLANDO, FL 32813

DR. ALFRED F. SMODE
TRAINING ANALYSIS AND EVALUATION
NAVAL TRAINING EQUIPMENT CENTER
CODE N-OOT
ORLANDO, FL 32813

DR. GARY POOCK
OPERATIONS RESEARCH DEPARTMENT
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93940

DR. A. L. SLAFKOSKY
SCIENTIFIC ADVISOR
COMMANDANT OF THE MARINE CORPS
CODE RD-1
WASHINGTON, D.C. 20380

DR. JOSEPH ZEIDNER
ACTING TECHNICAL DIRECTOR
U.S. ARMY RESEARCH INST.
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

TECHNICAL DIRECTOR
U.S. ARMY HUMAN ENGINEERING LABS
ABERDEEN PROVING GROUND
ABERDEEN, MD 21005

U.S. AIR FORCE OFFICE OF
SCIENTIFIC RESEARCH
LIFE SCIENCES DIRECTORATE, NL
BOLLING AIR FORCE BASE
WASHINGTON, D.C. 20332

DR. DONALD A. TOPMILLER
SYSTEMS ENGINEERING BRANCH
HUMAN ENGINEERING DIVISION
USAF AMRL/HES
WRIGHT-PATTERSON AFB, OH 45433

LT. COL. JOSEPH A. BIRT
HUMAN ENGINEERING DIVISION
AEROSPACE MEDICAL RESEARCH LAB.
WRIGHT PATTERSON AFB., OH 45433

AIR UNIVERSITY LIBRARY
MAXWELL AIR FORCE BASE, AL 36112

DR. ROBERT WILLIGES
HUMAN FACTORS LABORATORY
VIRGINIA POLYTECHNIC INSTITUTE
130 WHITTEMORE HALL
BLACKSBURG, VA 24061

DR. ARTHUR I. SIEGEL
APPLID PSYCHOLOGICAL SERVICES
404 EAST LANCASTER STREET
WAYNE, PA 19087

DR. ROBERT R. MACKIE
HUMAN FACTORS RESEARCH, INC.
SANTA BARBARA RESEARCH PARK
6780 CORTONA DRIVE
GOLETA, CA 93017

DR. GERSHON WELTMAN
PERCEPTRONICS, INC.
6271 VARIEL AVENUE
WOODLAND HILLS, CA 91364

DR. J. A. SWETS
BOLT,BERANEK & NEWMAN, INC
50 MOULTON STREET
CAMBRIDGE, MA 02138

DR. ROSS L. PEPPER
NAVAL OCEAN SYSTEMS CENTER
HAWAII LABORATORY
P.O. BOX 997
KAILUA, HI 96734

DR. MEREDITH CRAWFORD
5606 MONTGOMERY STREET
CHEVY CHASE, MD 20015

DR. ROBERT G. PACHELLA
UNIVERSITY OF MICHIGAN
DEPT. OF PSYCHOLOGY
HUMAN PERFORMANCE CENTER
330 PACKARD ROAD
ANN ARBOR, MI 48104

DR. ROBERT FOX
VANDERBILT UNIVERSITY
DEPARTMENT OF PSYCHOLOGY
NASHVILLE, TN 37240

DR. JESSE ORLANSKY
INSTITUTE FOR DEFENSE ANALYSES
400 ARMY-NAVY DRIVE
ARLINGTON, VA 22202

DR. STANLEY DEUTSCH
OFFICE OF LIFE SCIENCES
HQS, NASA
600 INDEPENDENCE AVE.
WASHINGTON, D.C. 20546

DR. WILLIAM A. MC CLELLAND
HUMAN RESOURCES RESEARCH OFFICE
300 N. WASHINGTON STREET
ALEXANDRIA, VA 22314

DR. WILLIAM R. UTTAL
UNIVERSITY OF MICHIGAN
INSTITUTE FOR SOCIAL RESEARCH
ANN ARBOR, MI 48106

DR. DAVID J. GETTY
BOLT, BERANEK & NEWMAN
50 MOULTON STREET
CAMBRIDGE, MA 02138

DR. RICHARD R. ROSINSKI
UNIVERSITY OF PITTSBURGH
DEPARTMENT OF INFORMATION
SCIENCE
PITTSBURGH, PA 15260

DIRECTOR, HUMAN FACTORS WING
DEFENSE & CIVIL INSTITUTE OF
ENVIRONMENTAL MEDICINE
P.O. BOX 2000
DOWNSVILLE, TORONTO, ONTARIO
CANADA

DR. A.D. BADDELEY
DIRECTOR, APPLIED PSYCH. UNIT
MEDICAL RESEARCH COUNCIL
15 CHAUCER ROAD
CAMBRIDGE, CB2 2EF
ENGLAND