Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

by Dominic W. Massaro and Erika Ferguson

Category width, as defined by Pettigrew's (1958) Category Width Scale, is a cognitive variable that purportedly reveals individual differences in categorization strategy. Subjects differ in terms of broadness and narrowness of judgments of category width--to what extent they will accept exemplars as good instances of a category. We tested the hypothesis that category width would be related to how subjects behave in different speech perception tasks. Differences found between extremely broad and narrow categorizers on such tasks would be helpful in understanding the nature of the perceptual and cognitive processes underlying the category width distinction. No effects attributable to category width were found when results were analyzed in terms of subjects' (a) discrimination and feature evaluation of auditory and visual information in speech events, (b) integration of these sources of information, (c) the process of decision, and (d) subjective preference for a two-choice versus a nine-choice response method. The results from both male and female and broad and narrow categorizers supported the predictions made by a fuzzy logical model of perception (FLMP). In the FLMP, people have access to continuous information about each feature of a stimulus, they make independent evaluations of each feature based on this information, the various features are integrated, and a decision is made based on the relative support for the viable alternatives. Given the common processes involved in speech and other pattern perceptual-recognition tasks, we conclude that fundamental processes involved in pattern recognition are unlikely to vary with personality measures, such as category width.

Categorization has proved to be central to many areas of performance and skill (Medin & Smith, 1984; Mervis & Rosch, 1981; Weiskrantz, 1985). Categorization has functional value in that it allows us to treat different things as if they were identical. Thus, we can act on an object we have never seen before as appropriate for its category membership. Even the development of expertise has been interpreted in terms of developing highly efficient and informative categories. Experienced students categorize physics problems according to theoretical principles, whereas novices categorize on the basis of superficial surface features (Larkin, McDermott, Simon, & Simon, 1980). Understanding the principles of categorization, therefore, offers the potential of fundamental insights into complex cognitive behavior.

Progress in making transparent the principles of categorization accelerated significantly with the seminal work of Rosch and her colleagues (Rosch, 1978; Rosch & Mervis, 1975). In several experiments, Rosch and others have shown that not all exemplars are equally good members of a category. Some are highly prototypical, others are only reasonable instances, and still others are peripheral or borderline cases of category membership. Contrary to the classical theory of categories, its exemplars cannot be defined in terms of necessary and sufficient properties. A more useful metaphor is that of family resemblance as discussed by Wittgenstein (1953). Exemplars of a category bear a family resemblance to one another, although they do not necessarily have a set of common properties.

The importance of categorization and the recent progress in understanding it might also inform social and personality issues (Cantor & Mischel, 1979). When the variety of personality variables was considered, it became apparent that the concept of cognitive style might be relevant to categorization research. One measure of cognitive style of some interest during the last 3 decades has been category width: the range of instances included in a cognitive category (Pettigrew, 1982). Individuals vary significantly in regard to category widths. Some will accept exemplars of the category "table" only if they are ideal instances with the appropriate attributes (flat top, reasonable size, and four legs of appropriate length). Others will accept less ideal instances (curved top and no legs). (The concept of category width not only is consistent with the fuzzy view of categories, it actually anticipated the findings of the more recent categorization research described in the previous paragraph. The individual differences imply a flexibility in categorization not acknowledged at that time in the cognitive-experimental literature.)

Category width is usually measured by Pettigrew's (1958) 20-question test. Each question states the average value of a property for members of a given category and asks for the largest and smallest values possible for members of that category. Four possible answers are given for each property. The categories chosen for the test were purposely obscure

Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

so that the answers would reflect a subject's general processing strategy and not situation-specific knowledge (we will have reason to question this assumption in the discussion section). For example, Question 1 (Pettigrew, 1958, p. 534) reads as follows:

It has been estimated that the average width of windows is 34 inches. What do you think:

a. is the width of the widest window . . .

1. 1,363 inches 2. 341 inches 3. 48 inches 4. 81 inches

b. is the width of the narrowest window . . .

1. 3 inches 2. 18 inches 3. 11 inches 4. 1 inch

The four possible answers to each were determined by giving pilot subjects an open-ended form of the test, and taking the 10th, 35th, 65th, and 90th percentile answers. The answer choice is scored in terms of how extreme it is. The definition of a broad categorizer, then, is one who tends to give extreme values, resulting (arbitrarily) in a high score. A narrow categorizer is one who lends to restrict his or her answers to a small range of values. These strategies are manifested when the actual values of category members (e.g., the true width of the largest or smallest existing window) are unknown to the subjects. If they were known, subjects could simply indicate the correct answers.

There have been other tests that profess to measure category width, such as object sorting tasks and other measures of "equivalence range." Pettigrew's measure, though, has been cited as being "a more pure index of category breadth" than other tests (Wallach & Kogan, 1965, p. 99).

The Category Width Scale has been given to thousands of subjects, with the conclusion that individuals vary with respect to the range of instances they include in a cognitive category (Pettigrew, 1982). Most research correlates category width with other variables to further characterize "broad" and "narrow" categorizers. Two of the most consistent correlations have been with sex (males lend to be broader categorizers than females) and with mathematical skill (those better in math tend to be broader). Others have proposed that category width is correlated with risk taking and creativity. There has been no clear explanation, however, for the individual difference being measured.

Our proposal is that insights into both category width and categorization can be made by providing a process-analysis of categorization behavior. By directly assessing psychological processes involved in categorization, we can understand what role, if any, category width might play. In addition, given the fundamental similarity between perceptual and conceptual categorization, these issues can be studied in the perceptual domain of speech. In fact, the original task after which Pettigrew modeled the Category Width Scale asked for subjects' judgments about perceptually oriented categories, such as the range of pitch of a female singing voice (Bruner & Rodriguez, cited in Bruner, Goodnow, & Austin, 1956).

The theoretical framework that we use for a process-analysis of categorization is a model that has been tested favorably in many domains (Massaro & Oden, 1980; Oden, 1981). The fuzzy logical model of perception (FLMP) has three processing stages called feature evaluation, feature integration, and decision (Massaro, 1987b, 1989). The FLMP is grounded on the premise of continuous information supporting categorization behavior. At the feature evaluation stage, all available sources of information, called features, are evaluated. During feature integration, prototype descriptions of the relevant categories are used to integrate or combine these features. The outcome gives a goodness-of-match value of the stimulus to each relevant category alternative. The decision stage categorizes the stimulus in terms of the relative goodness of match of the stimulus with the possible alternatives.

The present experiment also allows a further comparative examination of the FLMP with the categorical model of perception (CMP). The CMP assumes that only categorical information is available from the auditory and visual sources and that the identification judgment is based on separate decisions to the auditory and to the visual sources (Massaro & Cohen, 1983; McGurk & MacDonald, 1976). Although the FLMP has been shown to give a superior description of performance relative to the CMP, the current study broadens this comparison in several ways. First, the models will be tested against subjects varying in both category width and sex. It is possible that the relative goodness-of-fit of the models differs as a function of these individual differences. Second, previous comparisons have usually been limited to identification tasks (e.g., Massaro & Cohen, 1983). The comparisons in the present study will be extended to include a nine-interval rating task, a same-different discrimination task, and the relationship between two-choice identification and discrimination. Both the FLMP and the CMP will be formalized to give quantitative predictions of these results. Thus, this study offers informative tests of the FLMP and CMP independently of the outcome of the individual-differences analysis.

The domain of this study is bimodal speech perception in which subjects process speech events consisting of both auditory and visual information. The task is a replication and extension of the Massaro and Cohen (1983) study to evaluate the role of category width, as measured by the Pettigrew scale. Subjects categorize speech events seen on a videotape as being either the syllable /ba/ or /da/ (/a/ as in "father"). The actual events were dubbed composites of one of three possible visual stimuli (a man articulating /ba/, /da/, or not moving his lips at all) paired with one of nine possible auditory stimuli (synthesized computer speech sounds consisting of nine equally spaced intervals on a continuum between and including /ba/ and /da/). Subjects performing this task are influenced not only by the auditory source of information but also tend to be influenced by the visual information even when instructed to indicate only what they heard. In addition, subjects are increasingly influenced by the visual source as the ambiguity of the auditory speech increases.

It might appear to some that speech perception is an unlikely domain in which to find individual differences. However, our previous research has shown that the processes involved in speech perception are highly representative of pattern recognition and categorization in a variety of other domains (Massaro, 1987b). Speech perception is representative of other perceptual and cognitive tasks, including depth perception, location perception, object recognition, reading, sentence interpretation, memory retrieval, person impression, and decision making. Each of these domains makes available multiple sources of continuous information, and all of these sources are evaluated and integrated to achieve perception, cognition, and action. Therefore, speech perception is a reasonable domain to explore the contribution of individual differences to performance.

In terms of the FLMP, differences in category width might be reflected in any of the three processes involved in categorization. Feature evaluation might account for observed differences in category width. Differences at this stage could come about for a number of reasons. First, subjects with different category widths might differ in the discrimination of the various properties of the speech syllables. Second, the relative contribution of the visual and auditory sources might differ for broad and narrow categories. Third, the visual and auditory sources of information might be evaluated independently of one another for one group of subjects but not for the other. Fourth, the feature evaluation might provide continuous information for broad categorizers, for example, and only discrete or categorical information for narrow categorizers. Any of these four differences would be apparent in the results and the tests of the FLMP and CMP.

Differences in category width could also result from differences in the integration of the auditory and visual features of the speech syllable. Various forms of integration are possible and these would be revealed in the present study. The FLMP assumes a multiplicative integration in which the least ambiguous source of information has the greatest impact on performance. Narrow categorizers, on the other hand, might average the various sources which would tend to give more restricted categories. Finally, decision might differ for the broad and narrow categorizers. Such a difference should also be revealed in the analysis of the experimental results.

In addition to these two tasks, a discrimination task was used to assess how well subjects distinguish between various stimuli. Subjects were required to answer same or different to pairs of the test stimuli. Two of the auditory levels from the identification part of the experiment, one from the /ba/ end and one from the /da/ end of the continuum, were combined with the three visual levels, /ba/, /da/, and neutral. We might expect broad categorizers to be biased to say same relative to narrow categorizers, because they should adopt a broader category for same. Group differences found between broad and narrow categorizers on the discrimination task could also lend support to the notion of perceptual processes being a major contributing factor to the category width distinction.

Another possible difference between broad and narrow categorizers is one of subjective preference for certain methods of relating to the world. "To test this possibility, all subjects were given a questionnaire comparing their attitudes toward each of the methods of categorizing the syllables used in the first two parts of the experiment, the two-choice (identification) and the nine-choice (rating) method. Several notions of category width would predict that narrow categorizers would prefer the nine-choice continuum, whereas broad categorizers would prefer the dichotomous identification method. For

Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

instance, in the realm of creativity, divergent thinking tends to be more characteristic of broad categorizers. Divergent thinking leads to innovative and useful solutions to a problem rather than one correct answer and also demands a more holistic strategy of approach (Mayer, 1983). Thus broad categorizers who use divergent thinking may prefer a two-choice task, and should not be bothered by maintaining broad categories given only two choices. The narrow categorizers, who think more convergently, would be expected to prefer the nine-choice task, because they could maintain narrower categories in their judgments. In terms of this study, then, broad categorizers might prefer a two-choice identification task, even if they show no differences from narrow categorizers when actually performing it.

The goal here is to investigate the relationship between the hypothesized cognitive strategy of category width and the component processes involved in speech perception. Speech perception exemplifies a categorization domain containing the fundamental processes of feature evaluation, feature integration, and decision (Massaro, 1989; Massaro & Oden, 1980). The use of a variety of processing tasks and testing process models of performance in studying differences between broad and narrow categorizers would help to clarify the nature of their processing differences. In addition, differences might illuminate some common controversies in the speech perception literature, such as whether information is perceived in a continuous or a categorical manner.

EXPERIMENT

METHOD

Subjects

We recruited 29 subjects (17 male, 12 female) from two introductory psychology classes. They were given a choice of extra course credit or $5.00 for their participation. Subjects selected had taken Pettigrew's Category Width Scale on the first day of class and scored among the top or bottom 10% for their sex in their respective classes. Selected subjects were unaware of why they in particular were being asked to participate. For the final data analysis, the 20 subjects with the fewest missed responses in the experimental tasks were used, such that the n for each cell in the Category Width x Sex analyses was 5.

Tests and stimuli

The videotape used in Parts 1 and 2 of the experiment was prepared identically to the stimuli used by Massaro and Cohen (1983, Experiment 2). On each trial the videotape showed a male speaker saying either /ba/ or /da/ or keeping his lips closed (the "neutral" condition). The auditory stimuli were synthesized from characteristics of the speaker's voice. They formed a continuum of nine sounds equally spaced along a continuum from /ba/ to /da/.

The Natural Category Width Scale consisted of eight questions, each of which listed a common semantic category and four possible members of the category. The questions asked subjects to indicate the instances they considered to be (or not to be) members of the given category. On Version 1 of the questionnaire, odd-numbered categories asked for inclusion of members, and even-numbered categories required exclusion of nonmembers from among the listed instances. Version 2 had the opposite format. The four instances listed for each semantic category were taken from a previous study that asked subjects to rate the extent an instance represented the rater's ideal meaning of the category (Rosch, 1975). This scale was scored in the same manner as the Pettigrew (1958) test. Each item was assigned the value 4, 3, 2, or 1 according to its "degree of membership" in the category. These membership values were derived from the ratings of Rosch's (1975) subjects of the extent to which each instance represented the ideal meaning of the category. An instance about 10% from the top of the list of instances ranked in terms of the ratings was given a value of 1. Instances 35%, 65%, and 90% from the top of the list were given values of 2, 3, and 4, respectively. Thus, as in the Pettigrew test, more extreme instances received higher values. Given eight test items, the maximum score is 80 and the minimum score zero.

The Preference Questionnaire consisted of five questions concerning the subjects' impressions of the two-choice versus the nine-choice response methods. The five questions asked which of the two response methods (a) was easiest to use, (b) seemed more accurate, (c) made you more uncomfortable, (d) made you more unsure of your answers, and (e) would you prefer if you had to do the task again. It also checked for subject bias by asking subjects what they believed was the

Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

purpose of the experiment.

Procedure

Three subjects were tested simultaneously in individual rooms. On each trial, one of the nine auditory stimuli on the continuum from /ba/ to /da/ was paired with one of the three possible visual stimuli, /ba/, /da/, or neutral, as synchronized by the computer according to the method described in Massaro and Cohen (1983). The stimuli were presented in 11 blocks of the 27 possible combinations plus 10 practice trials for a total of 307 trials. Subjects were instructed to "watch a speaker and listen to what is spoken" and simply to "indicate whether the man said /ba/ or /da/." They were warned that "on some trials you will see the speaker say the sound, and on other trials the sound will be presented but the speaker will not move his mouth." Subjects had 2,750 ms in which to respond by pressing either key B or D on a keyboard positioned to the side of the TV monitor.

Subjects were given the Natural Category Width Scale under the impression that it was unrelated to the task they had just performed. Then subjects were instructed that, because they had probably noticed that some of the trials in the first session had been difficult to classify as either /ba/ or /da/, they were going to have a chance to make finer distinctions among the stimuli. They were to "identify what the man said on a continuum from 1 to 9 with 1 being the best possible /ba/ and 9 being the best possible /da/." The stimuli and testing conditions were identical to those used in the two-choice task except that subjects responded to each trial by pressing a number from 1 to 9 on the keyboard. After this session, subjects completed the Preference Questionnaire comparing the two methods of responding.

On each trial in the discrimination task, a pair of stimuli was presented. For each member of the pair, one of the three visual levels was synchronized with either Auditory Level 3 or 7 from the auditory continuum. Table 1 shows the groupings of trial types for this part of the experiment. The four trial types were same, auditory different, visual different, and both different. There were actually 70% different trials and 30% same trials. The stimulus pairs were presented in 10 blocks of the 20 possible combinations described in Table 1. About 1,250 ms occurred between the onsets of each speech TABULAR DATA OMITTED event, and subjects had about 2 s to make a response. Subjects were instructed to respond either same or different to each pair of stimuli by pressing one of two keys on the keyboard. They were told to answer same only if they thought both members of the pair were identical.

At the close of the experiment subjects were debriefed, at which time we explained the significance of the category width scales to the tasks they had just performed and to how they had been chosen as subjects.

RESULTS

Category width

An analysis of variance (ANOVA) was carried out on the category width scores with category width and sex as factors. The average category width score for the 10 broad categories was 93.2, significantly higher (broader) than the corresponding average of 53.2 for the narrow categories, F(1, 16) = 67.8, p |is less than~ .001. Consistent with previous findings, males had higher scores (79.8) than females (66.6), F(1, 16) = 7.39, p |is less than~ .025, within both the broad and narrow groups. That is, there was no interaction of category width and sex. Thus, the factorial design is successful in providing an evaluation of category width without a confounding by sex differences.

Choice responses

The two-choice identification and nine-choice rating tasks were analyzed using an ANOVA. The proportion of /da/ responses was used as the dependent measure in the two-choice task and the degree of /da/ rating was the dependent measure in the nine-choice task. The latter rating was computed as a value between 0 and 1 by treating the 1 through 9 responses as an interval scale with 1 and 9 corresponding to the 0 and 1 value of the degree of /da/, respectively. An ANOVA was carried out on P(/da/) and ratings, with category width, sex, auditory level, and visual level factors.

Figure 1 gives the average results for the two-choice tasks. A comparison of the left and right panels shows no significant effect of category width nor interaction with the other variables (all Fs ranging from 1 to 16 df in the numerator and from 16

Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

to 256 df in the denominator were less than 1). Figure 2 shows that the two groups differing in category width also behaved similarly to one another in the nine-choice rating task. There was no interaction between category width and method of response. In addition, there was no significant difference due to sex on either task, and sex did not interact with any of the other variables.

As can be seen in Figures 1 and 2, the proportion of /da/ responses tended to increase as the syllables changed from the /ba/ end to the /da/ end of the auditory continuum. For the most /ba/-like to the most /da/-like stimuli, these proportions ranged from .129 to .891, F(8, 128) = 139.5, p |is less than~ .01, and .236 to .782, F(8, 128) = 39.66, p |is less than~ .01, with the two-choice and nine-choice methods of response, respectively. There was also a main effect along the visual continuum, with the /ba/ articulation eliciting the smallest proportion of /da/ responses with both the two-choice (.334) and the nine-choice (.354) methods, /da/ the largest (.755 and .698), and the neutral trials an intermediate proportion (.604 and .576), Fs(2, 32) = 51.11 and 52.04, p |is less than~ .025. There was also a significant interaction between the visual and auditory variables, Fs(16, 256) = 13.97 and 15.95, p |is less than~ .01, with the visual variable having its greatest effect in the most ambiguous region of the auditory continuum.

Reaction times

Figure 3 gives the average reaction times (RTs) for the two-choice task as a function of the auditory and visual levels. Reaction times changed systematically with changes in the bimodal speech event, F(16, 256) = 21.2, p |is less than~ .001. With no articulation, RTs were longer to sounds toward the middle of the auditory continuum. With a /ba/ articulation, RTs increased from the /ba/ to the /da/ end of the auditory continuum, and the opposite was the case for the /da/ articulation. These results replicate exactly previous studies (Massaro & Cohen, 1983) which were best explained in terms of the RT increasing with increases in the overall ambiguity of the speech event (Massaro, 1987b, chap. 4). No significant differences due to category width or sex were found, and these variables did not interact with each other or with the stimulus variables.

Natural Category Width Scale

Given eight test items on this scale, the maximum score is 80 and the minimum score is zero. The overall average score of 57.8 shows that the test would have been in a sensitive range to observe individual differences if they existed. There was no significant difference due to sex (males = 57.4; females = 58.2), and broad categorizers had only a slightly larger score (60.4) relative to narrow categorizers (55.2), F(1, 16) = 5.63, p |is less than~ .05.

Subjects' scores on the two category-width tests were not significantly correlated (r = .268, p |is greater than~ .3). The natural scale was designed to overcome the main disadvantage of Pettigrew's scale--its reliance on quantitative judgments (Pettigrew, 1982), and thus the lack of a significant correlation is disappointing because the two scales should be measuring the same cognitive trait.

Preference Questionnaire

The Preference Questionnaire revealed a marginal sex difference in that more females than males found the nine-choice task easier, F(1, 16) = 3.20, p |is less than~ .10. All subject groups perceived the nine-choice method to be more accurate, with no significant category width or sex differences. There was no difference on the questions concerning discomfort with answer method between broad and narrow categorizers. No effect of category width was found on subjects' preference for one of the two methods of responding. Table 2 shows that a comparable percentage of subjects preferred each of the methods. On the question, "With which answer method were you least sure of your responses?" a marginal sex difference was found, F(1, 16) = 3.20, p |is less than~ .10. Females answered two more often than nine.

Two- and nine-choice model tests

The FLMP and the CMP were fit to the identification and rating results. These models are described in Massaro and Cohen (1983). When the FLMP is applied to the present task using auditory and visual speech, both sources are assumed to provide continuous and independent evidence for the alternatives /ba/ and /da/. Defining the onsets of the second (F2) and third (F3) formants as the important auditory feature and the degree of initial opening of the lips as the

Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

important visual feature, the prototype for /da/ would be

/da/ : Slightly falling F2-F3 and open lips.

The prototype for /ba/ would be defined in an analogous fashion:

/ba/ : Rising F2-F3 and closed lips.

Given a prototype's independent specifications for the auditory and visual sources, the value of one source cannot change the value of the other source at the prototype-matching stage. The integration of the features defining each prototype is determined by the product of the feature values (see Massaro, 1987b, chap. 7, for supporting evidence). If |a.sub.i~ represents the degree to which the auditory stimulus |A.sub.i~ supports the alternative /da/, that is, has slightly falling F2-F3, and |v.sub.j~ represents the degree to which the visual stimulus |V.sub.j~ supports the alternative /ba/, that is, has open lips, then the outcome of prototype matching for /da/ would be

TABULAR DATA OMITTED

/da/ : |a.sub.i~|v.sub.j~,

where the subscripts i and j index the levels of the auditory and visual modalities, respectively. With two response alternatives, the auditory support for the alternative /ba/ is simply 1 - |a.sub.i~, and the visual support is 1 - |v.sub.j~ (Massaro, 1989). Thus, the outcome of prototype matching for /ba/ would be

/ba/ : (1 - |a.sub.i~)(1 - |v.sub.j~).

The decision would determine their relative merit, leading to the prediction that

P(/da/ / |A.sub.i~|V.sub.j~) = |a.sub.i~|v.sub.j~ / |summation~

where |sigma~ is equal to the sum of the merit of the /ba/ and /da/ alternatives. Equation 1 also gives the prediction for the nine-choice task when the response is treated as a rating between 0 and 1 (as we did in the analysis of the results).

The CMP assumes that the auditory source of information provides only categorical information, as does the visual. Thus, the identification judgment is based on categorical decisions to the auditory and visual sources. Considering the /ba/ identification, the visual and auditory decisions could be /ba/-/ba/, /ba/-/da/, /da/-/ba/, or /da/-/da/. If the two decisions to a given speech event agree, the identification response can follow either source. When the two decisions disagree, it is assumed that the subject will respond with the decision to the auditory source on some proportion p of the trials, and with the decision to the visual source on the remainder (1 - p) of the trials. The weight p reflects the relative dominance of the auditory source.

The probability of a /da/ identification response, P(/da/), given a particular auditory/visual speech event, |A.sub.i~|V.sub.j~, would be

|Mathematical Expression Omitted~

where i and j index the levels of the auditory and visual modalities, respectively. The |a.sub.i~ value represents the probability of a /da/ decision given the auditory level i, and |v.sub.j~ is the probability of a /da/ decision given the visual level j. The value p reflects the bias to follow the auditory source. Each of the four terms in the equation represents the likelihood of one of the four possible outcomes multiplied by the probability of a /da/ identification response given that outcome. To fit this model to the results, each unique level of the auditory stimulus requires a unique parameter |a.sub.i~, and analogously for |v.sub.j~.

The quantitative predictions of each model are determined by using the parameter-estimation program STEPIT (Chandler, 1969). The model is represented to the program in terms of a set of prediction equations and a set of unknown

Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

parameters. By iteratively adjusting the parameters of the model, the program minimizes the squared deviations between the observed and predicted points. The outcome of the STEPIT program is a set of parameter values which, when put into the model, come closest to predicting the observed results. For the FLMP, for example, values for |a.sub.i~ and |v.sub.j~ are determined reflecting the outcome of the evaluation operation that gives the degree to which a given level of a given source of information supports the alternative /da/. Thus, STEPIT maximizes the accuracy of the description of each model. The goodness-of-fit of the model is given by the root mean square deviation (RMSD)--the square root of the average squared deviation between the predicted and observed values.

To predict the 27 stimulus conditions, the FLMP requires 3 + 9 = 12 parameters, whereas the categorical model requires 3 + 9 + 1 = 13 parameters. The dependent measure was the proportion of /da/ identifications in the two-choice task and the degree of /da/ rating in the nine-choice task. The models were fit to the results of each subject individually and to the average results of each of the four groups of subjects.

Table 3 gives the average RMSD values for the fit of the individual subjects and the RMSD values for the average subject. The FLMP gives a consistently better description of the results for all subjects, TABULAR DATA OMITTED regardless of category width or sex. Figure 4 gives the predictions of the FLMP, and Figure 5 gives the predictions of the CMP. As can be seen in the figures, the contribution of the visual source is greatest when the auditory source is ambiguous. The FLMP captures this result, whereas the CMP does not.

Table 3. Root mean square deviations (RMSDs) for subjects using the two-choice and nine-choice answer methods with the fuzzy logical model of perception (FLMP) and the categorical model (CMP)

Categorizer	Two-choice		Nine-choice	
group	FLMP	CMP	FLMP	CMP
Male/broad	.0352	.1167	.0431	.0481
X	.0322	.1677	.0446	.0955
Female/broad	.0232	.1054	.0356	.0514
X	.0319	.1677	.0446	.0955
Male/narrow	.0345	.1007	.0350	.0497
X	.0347	.1582	.0447	.0678
Female/narrow	.0276	.1211	.0373	.0800
X	.0282	.1578	.0506	.1076

Note. The first line for each categorizer group is the RMSD for the average of the subjects, and the second line (X) is the average RMSD for the fit of the individual subjects.

Discrimination task

Percentage correct discrimination was analyzed as a function of the four kinds of trials: same, auditory-different, visual-different, and both different. Category width and sex were included in the analysis. Figure 6 gives the average percentage of correct discrimination as a function of the four kinds of trials; subjects were least accurate under the auditory-only condition, F(3, 48) = 22.22, p |is less than~ .001. However, on the discrimination task, as in the identification, there were no main effects of category width or interactions with trial type or sex, all Fs |is less than~ 1.

Identification/discrimination model tests

The FLMP and CMP were contrasted in three separate tests against the identification and discrimination tasks. The three tests differed with respect to the number of free parameters used in the models and the results that were predicted. In the first contrast, the identification results were used to predict the discrimination results, as is standard in the identification/discrimination task. These tests used the identification results to predict discrimination, and thus no free parameters were used in the test. In the second contrast, free parameters were estimated to predict just the discrimination results. In the third contrast, free parameters were estimated to predict both identification and discrimination.

There were 20 unique trials in the discrimination task, as can be seen in Table 1, and 6 relevant identification trials from the two-choice task. In the first contrast, the 6 identification probabilities were used to predict discrimination for the

Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

categorical model (see Massaro, 1987, chap. 4, for a more complete description of this type of analysis). For the FLMP, the a_i and v_j parameter values that were estimated to predict identification were used to predict discrimination. Based on the derivation for disjunction for the FLMP, an auditory difference value d_a and a visual difference value d_v were computed by subtracting the two relevant truth values

$$d_a = a_7 - a_3$$

$$d_v = v_D - v_B,$$

where a_7 and a_3 are the auditory degree of /da/-ness for Auditory Levels 7 and 3 and v_D and v_B are the visual degree of /da/-ness for a /da/ and /ba/ visual articulation, respectively.

In the second contrast, six identification probabilities were estimated as free parameters to test the categorical model, and two difference values (d_a and d_v) were estimated to test the FLMP. This test allows for some change in the relevant parameters for the two models from the identification to the discrimination task. In the third contrast, six identification probabilities were estimated to test the categorical model simultaneously against both identification and discrimination. For the test of the FLMP, five parameters had to be estimated (a_7, a_3, v_D, v_B, and v_N) to predict both identification and discrimination.

The RMSD between predicted and observed values provides a measure of the accuracy of the predictions. The RMSD values for the three contrasts are given in Tables 4, 5, and 6. These values are based on the fit of the average performance of each of the four groups of subjects. The parameter values are given in Tables 7 and 8. As can be seen in the RMSD values, the description given by the FLMP was superior in every contrast to that given by the categorical model. This result is particularly impressive in Contrast 2 in which the categorical model requires three times as many parameters as the FLMP and yet the RMSD for the FLMP was less than half of that given by the categorical model. Thus, the FLMP gives a much better description of performance of the identification/discrimination task in the same domain traditionally used to support categorical perception.

Table 4. Root mean square deviations (RMSDs) for the categorical model (CMP) and the fuzzy logical model of perception (FLMP) for predicted discrimination based on the identification results (Contrast 1)

Categorizer group	Model CMP	FLMP
Male/broad	.252	.153
Male/narrow	.214	.131
Female/broad	.222	.129
Female/narrow	.243	.164
Average	.232	.144

Table 5. Root mean square deviations (RMSDs) for the categorical model (CMP) and the fuzzy logical model of perception (FLMP) for predicted discrimination based on the estimation of free parameters (Contrast 2)

Categorizer group	Model CMP	FLMP
Male/broad	.197	.104
Male/narrow	.195	.123
Female/broad	.193	.078
Female/narrow	.184	.068
Average	.192	.093

There were no major differences for the four different groups in terms of the descriptions given by the models. The FLMP was consistently better for all groups in all of the contrasts. The only noticeable difference was a slightly better description given by the FLMP for the females in Contrasts 2 and 3. Consistent with the negative findings in the identification and rating tasks, we found no evidence for differences in the processes involved in discrimination as a function of category

Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

width or sex.

Power

Of important concern in the present series of experiments is a potential lack of power--a Type I error of failing to reject the null hypothesis when it is, in fact, false. However, the experiment was powerful enough to find significant main effects of the auditory and visual independent variables, and an interaction between them. Even so, it is possible that broad and narrow categorizers do process speech differently from one another, but we simply failed to detect this difference because of low power. We adopted the strategy of testing a few subjects in each group intensively, rather than testing many subjects for just a few observations. Our strategy should increase power because of the increased fidelity that is achieved in intensive testing with many repeated measures on the same subject. In fact, we actually eliminated 9 of the 29 subjects in the analysis to keep an equal number of subjects per cell. The conclusions did not change when we included these additional 9 subjects in the analysis.

Table 6. Root mean square deviations (RMSDs) for the categorical model (CMP) and the fuzzy logical model of perception (FLMP) for both predicted identification and discrimination based on the estimation of free parameters (Contrast 3)

	Model	
Categorizer group	CMP	FLMP
Male/broad	.189	.104
Male/narrow	.177	.110
Female/broad	.179	.075
Female/narrow	.178	.071
Average	.180	.090

TABULAR DATA OMITTED

To address the power issue, we asked whether the present experiment would have been powerful enough to detect a difference between broad and narrow categorizers--if in fact a difference existed. In one hypothetical analysis, it was assumed that the broad categorizers TABULAR DATA OMITTED gave results in line with the CMP, whereas the narrow categorizers gave results predicted by the FLMP. Ten hypothetical broad categorizers were created by generating hypothetical results using the parameter values from the fit of the CMP to the actual results of the broad categorizers. Variability was added to these ideal results by adding the variability that we actually observed in our experiment. The M|S.sub.e~s for these two nonsignificant Group x ISI interactions were used as the estimates of variability. Ten hypothetical narrow categorizers were generated in a similar manner by taking the parameter values from the fit of the FLMP to the actual results of the 10 narrow categorizers. The 10 hypothetical broad categorizers were contrasted with the 10 hypothetical narrow categorizers in an ANOVA. The results revealed a highly significant three-way interaction of category width, auditory, and visual factors, $F(16, 256) = 3.48$ and 3.38 for the two-choice and nine-choice tasks, respectively. Thus, our experiment was clearly powerful enough to find a significant difference of the order of the difference predicted by the FLMP and CMP; it is unlikely that our null results were due to low power.

DISCUSSION

One explanation for the lack of significant findings in this study is simply that whatever is causing the consistent individual differences on the Category Width Scale is not people's perceptual categorization process. The reason may be that the processes of speech perception, as required in the tasks presented here, are far removed from the cognitive strategy of category width. This could also explain the contradictory results of the few previous studies on category width that were perceptually oriented. For instance, narrow categorizers were shown to possess greater accuracy in judging the relative size of circles (Huckabee, 1976), but broad categorizers were superior in judging the lengths of lines (Bieri, 1953). It is difficult to imagine that these two tasks are fundamentally different, and, therefore, the different results are probably due to statistical error. It seems safest to conclude that the fundamental processes involved in perceptual recognition are independent of test differences on Pettigrew's (1958) category width questionnaire. To what extent the questionnaire

Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

performance is related to other aspects of cognitive functioning remains an open question.

It would be easy enough, especially for believers in category width, to discard the current results simply because they are negative. This action would represent a confirmation bias (Popper, 1959), however, creating a nonproductive research strategy. A recent review of confirmation bias in the study of the sleeper effect in attitude change reminds us of the gravity of such a bias (Greenwald, Leippe, Pratkanis, & Baumgardner, 1986). For advocates of the concept of category width, the present results should, at least, be considered as representing the limiting conditions for category width influences.

The failure to find category width effects might be taken to support the modularity hypothesis that speech is uniquely handled by a specialized speech system that would supposedly be impenetrable by cognitive variables such as category width (Fodor, 1983; Liberman & Mattingly, 1985). However, the processes that have been observed in the speech domain have been observed in a variety of other perceptual and cognitive domains (Massaro, 1987a, 1987b). The processes assumed by the FLMP have found empirical support in reading, visual perception of depth, sentence interpretation, decision making, category judgments, person impression, and retrieval from memory. Thus, we expect that category width would also play very little role in these other domains.

The failure to find any influence of category width on pattern recognition warrants some concern about the validity of this concept as a description of perceptual/cognitive style. We also found little validity for the concept using a Natural Category Width test in which subjects indicated whether instances were considered to be members of a given category. Performance on the Natural Category Width test did not correlate with the Pettigrew scale, and there was only a 6.5% difference between broad and narrow categorizers relative to the 25% given by the Pettigrew scale. Thus, performance on the Pettigrew scale might be unique to the situation in which the participant knows very little about the range of instances being tested. For example, subjects given the mean size of windows are asked to estimate the range of window sizes. Some subjects give significantly larger estimations than other subjects. These same subjects, however, do not appear to have broader categories for concepts such as furniture, birds, and so on. Thus, the Pettigrew scale probably has very little to do with category representation per se, but might simply represent a numerical bias in tasks requiring quantitative answers. A test of this idea might take the form of asking for a quantitative estimate in domains that have little to do with familiar categories. One example is attendance: The introductory psychology course has an enrollment of 300; what is the range of attendance throughout the year? One should find the same differences between broad and narrow categorizers on questions of this sort if the differences are primarily due to numerical biases.

Another conceptualization of the category width phenomenon might warrant consideration. By design, the category width questionnaire represents a situation involving extremely limited information. The choices that are made might therefore depend on some unique strategic solution, which has little relationship to typical situations with much more information for action (Gibson, 1979). To use visual perception as an analogy, the identification of a completely impoverished visual display might have little to say about how normal visual perception proceeds. Such a display might not engage any of the processes typically involved in visual perception. Similarly, the category width questionnaire might not engage the fundamental processes involved in perception and cognition, and thus the negative findings should not be unexpected. What the original questionnaire does reflect, therefore, remains as much of a puzzle as ever.

Appendix. Categories and instances listed from most representative to least representative of the meaning of the category

Category	Instances
1. Furniture	Dresser, bookcase, cupboard, refrigerator
2. Fruit	Pear, blueberry, cranberry, tomato
3. Vehicles	Bus, airplane, rowboat, skis
4. Weapons	Switchblade, bullet, rocket, rope
5. Vegetables	Broccoli, turnip, peppers, pickles
6. Birds	Canary, hummingbird, crane, turkey
7. Sports	Canoeing, gymnastics, Ping-Pong, camping
8. Clothing	Blouse, overcoat, tie, ring

Notes

This study was undertaken as a senior thesis by Erika Ferguson under the direction of Dominic Massaro. Erika Ferguson

Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

is now at the University of California, Santa Barbara. The research was supported in part by NINCDS Grant 20314 from the Public Health Service and Grant BNS-83-15192 from the National Science Foundation. We would like to thank Thomas Pettigrew for making the category width scores of the subjects available, Michael Cohen for his important contribution to the experiment and data analysis, and Thomas Pettigrew, Philip T. Smith, and Quentin Summerfield for their comments on an earlier version of the article.

Correspondence concerning this article should be addressed to Dominic W. Massaro, Program in Experimental Psychology, University of California, Santa Cruz, CA 95064. Received for publication July 11, 1991; revision received October 29, 1991.

References

Bieri, J. (1953). Category width as a measure of discrimination. Journal of Personality, 22, 214-233.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). A study of thinking. New York: Wiley.

Cantor, N., & Mischel, W. (1979). Prototypes on person perception. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 12, pp. 3-52). New York: Academic Press.

Chandler, J. P. (1969). Subroutine STEPIT--Finds local minima of a smooth function of several parameters. Behavioral Science, 14, 81-82.

Fodor, J. A. (1983). Modularity of mind. Cambridge, MA: Bradford Books/MIT Press.

Gibson, J. J. (1979). The ecological approach to visual perception. Boston: Houghton Mifflin.

Greenwald, A. G., Leippe, M. R., Pratkanis, A. R., & Baumgardner, M. H. (1986). Under what conditions does theory obstruct research progress? Psychological Review, 93, 216-229.

Huckabee, M. W. (1976). Category width, confidence, expectancies, and perceptual accuracy. Bulletin in the Psychonomic Society, 7, 19-21.

Larkin, J. J., McDermott, J., Simon, D., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. Science, 208, 1335-1342.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. Cognition, 21, 1-36.

Massaro, D. W. (1987a). Integrating multiple sources of information in listening and reading. In D. A. Allport, D. G. MacKay, W. Prinz, & E. Scheerer (Eds.), Language perception and production: Shared mechanisms in listening, speaking, reading and writing (pp. 111-129). London: Academic Press.

Massaro, D. W. (1987b). Speech perception by ear and eye: A paradigm for psychological inquiry. Hillsdale, NJ: Erlbaum.

Massaro, D. W. (1989). Multiple book review of speech perception by ear and eye: A paradigm for psychological inquiry. Behavioral and Brain Sciences, 12, 741-794.

Massaro, D. W., & Cohen, M. M. (1983). Evaluation and integration of visual and auditory information in speech perception. Journal of Experimental Psychology: Human Perception and Performance, 9, 753-771.

Massaro, D. W., & Oden, G. C. (1980). Speech perception: A framework for research and theory. In N. J. Lass (Ed.), Speech and language: Advances in basic research and practice (Vol. 3, pp. 129-165). New York: Academic Press.

Mayer, R. E. (1983). Thinking, problem solving, cognition. New York: Freeman.

Cognitive style and perception: the relationship between category width and speech perception, categorization, and discrimination.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. Nature, 264, 746-748.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. Annual Review of Psychology, 35, 113-138.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. Annual Review of Psychology, 32, 89-115.

Oden, G. C. (1981). Fuzzy propositional model of concept structure and use: A case study in object identification. In G. W. Lasker (Ed.), Applied systems research and cybernetics (pp. 2890-2897). Elmsford, NY: Pergamon Press.

Pettigrew, T. W. (1958). The measurement and correlates of category width as a cognitive variable. Journal of Personality, 26, 532-544.

Pettigrew, T. F. (1982). Cognitive style and social behavior: A review of category width. In L. Wheeler (Ed.), Review of personality and social psychology (Vol. 3, pp. 199-223). Beverly Hills, CA: Sage.

Popper, K. R. (1959). The logic of scientific discovery. New York: Basic Books.

Rogers, R. I. (1973). Category width and decision making in perception. Perceptual and Motor Skills, 37, 647-652.

Rosch, E. (1975). Cognitive representations of semantic categories. Journal of Experimental Psychology: General, 104, 192-233.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), Cognition and categorization (pp. 27-48). Hillsdale, NJ: Erlbaum.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7, 573-605.

Wallach, M. A., & Kogan, N. (1965). Modes of thinking in young children. New York: Holt, Rinehart & Winston.

Weiskrantz, L. (Ed.). (1985). Animal intelligence. Oxford: Clarendon Press.

Wittgenstein, L. (1953). Philosophical investigations. New York: Macmillan.