

# HARD and SOFT interacting dimensions: Differential effects of dual context on classification

ROBERT D. MELARA and LAWRENCE E. MARKS

John B. Pierce Foundation Laboratory and Yale University, New Haven, Connecticut

In this article, we extend Garner's speeded classification procedure to investigate processes underlying the interaction of the auditory dimensions pitch, loudness, and timbre. In the experiments reported here, subjects classified attributes on these three auditory dimensions. Our extended procedure, called *multiclass*, is based conceptually on our model of how such dimensions interact; the model explains the perception of attributes from an attended dimension through the action of contextual constraints created by variation along an unattended dimension. Two forms of context are present simultaneously in each multiclass task: *intraclass context*, variation along the unattended dimension that interferes with the classification of attributes, and *redundant context*, variation along the unattended dimension that enhances classification. We find that such dual-context situations reliably distinguish two kinds of interacting dimensions. Subjects classifying HARD dimensions, here pitch and timbre, resist the ill effects of intraclass context and reap gains from redundant context. Subjects classifying SOFT dimensions, here loudness, show interference because the attributes are veiled perceptually in dual context. These findings, we argue, demonstrate the power of the multiclass procedure and fit well our view that dimensional interaction entails processing both at the level of the stimulus whole and at the level of stimulus attributes.

During the early 1970s, Garner (1974; Garner & Felfoldy, 1970) developed a reaction time (RT) technique known as *speeded classification* to study how perceptual systems process pairs of psychological dimensions. Since then, numerous experiments have used Garner's paradigm as a method for uncovering various forms of dimensional interaction, including integrality (e.g., Garner & Felfoldy, 1970), configural interaction (e.g., Pomerantz, 1986; Pomerantz & Garner, 1973), and correspondence interaction (e.g., Melara, 1989a; Melara & O'Brien, 1987). In our laboratory, we have recently extended Garner's procedure to probe more microscopically the structure of these dimensional interactions. We call this new technique the *multiple classification*, or *multiclass*, paradigm. The multiclass procedure is linked conceptually to a model of dimensional interaction that we have proposed recently (Melara & Marks, in press). In this article, we describe the multiclass procedure and its conceptual underpinnings, we report a series of experiments in which we employ the procedure, and we discuss the significance of the experimental results for our model.

## The Traditional Speeded Classification Paradigm

On the basis of Garner's (1974) work, one may distinguish two ways that a given pair of dimensions are processed perceptually, namely, as *interacting* or *separable*. Among the set of converging operations (Garner, Hake, & Eriksen, 1956) that Garner used to discriminate either dimensional interaction or dimensional separability is his speeded classification procedure. In a recent investigation, Melara and Marks (in press) used this paradigm to study processing of auditory stimuli; that investigation illustrates well the sets of tasks that constitute the procedure. Subjects (in Experiment 1A) judged either pitch or loudness under each of five conditions. Figure 1 shows values of the four stimuli used.

In each of four *baseline* tasks, subjects discriminated values from one dimension at a constant value of the other dimension (e.g., Stimulus A vs. Stimulus C). Performance in the baseline tasks was compared with that in *correlated-dimensions* tasks and in *filtering* tasks. In each of four correlated-dimensions tasks, values on the relevant dimension of a stimulus were paired with a redundant value on the nominally irrelevant dimension (e.g., Stimulus A vs. Stimulus B). With pitch and loudness, correlated dimensions led to superior performance over baseline—that is, to a redundancy gain. Finally, in each of two filtering tasks, all four stimuli were included and subjects attempted to attend selectively to one dimension, ignoring variation along the irrelevant dimension. For example, if pitch was the relevant dimension, subjects made one response to A or D, and a different response to B

Preparation of this paper was supported by NIH Postdoctoral Fellowship F32 HD07137-02 to Robert Melara and by NIH Grant NS21326 and NSF Grant BNS84-20017 to Lawrence Marks. We wish to thank Charles Eriksen, Daniel Algom, Bonnie Potts, and three anonymous reviewers for insightful comments on an earlier draft of this article. Please refer all correspondence to Robert D. Melara, whose present address is: Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907.

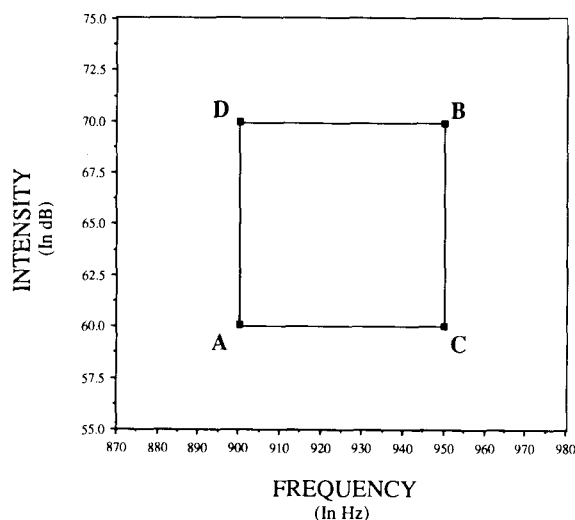


Figure 1. The values of the four stimuli used by Melara and Marks (in press) in their investigation of speeded classification of the dimensions pitch and loudness.

or C. A failure of selective attention is indicated by the presence of *Garner interference* (Pomerantz, 1983)—that is, impaired performance in filtering tasks relative to baseline tasks. With both pitch and loudness judgments, subjects displayed substantial Garner interference, indicating that these dimensions do, in fact, interact (see also Grau & Kemler Nelson, 1988; Wood, 1975).

#### HARD and SOFT Interacting Dimensions

In our conception, Garner's paradigm separately imposes two different and independent kinds of stimulus context on classification. These two contexts act to constrain or weight how perceivers process individual attributes on

dimensions. (Details regarding our model of contextual constraint will be reviewed in the General Discussion; see also Melara & Marks, in press). Varying dimensions orthogonally creates a stimulus context in which any given attribute may be perceived as itself varying; because that attribute must always receive the same response, this intraclass variability hinders performance. We call this first constraint *intraclass context*. Correlating dimensions creates a context in which two attributes are constrained to be more different from each other; because the two attributes are to be discriminated from each other, this greater perceptual distance facilitates performance. We call this second constraint *redundant context*. Our distinction between intraclass and redundant context forms the basis of the multiclass paradigm.

Consider Figure 2. In panel A we have schematized the typical Garner filtering condition: Subjects classify two values on Dimension X in the context of intraclass variability produced by Dimension Y. One can imagine shifting the two categories in opposite directions relative to the vertical axis, while maintaining a given level of intraclass variability. As one can see in panel B, such a shift would produce a correlation between the dimensions (redundant context) in the presence of intraclass context. Such *dual-context* classification is what we call "multiclass."

The basic idea behind the multiclass paradigm is simple: How is classification of attributes affected when two contexts—namely, redundant and intraclass—are present together in a single task? Subjects are instructed to classify attributes. Intraclass variability is present within each category but the dimensions are correlated. Multiclass thus combines the classification component of the filtering task with the redundancy of the correlated-dimensions task. We assume that these two contexts can operate independently on perception and so can produce demonstrably

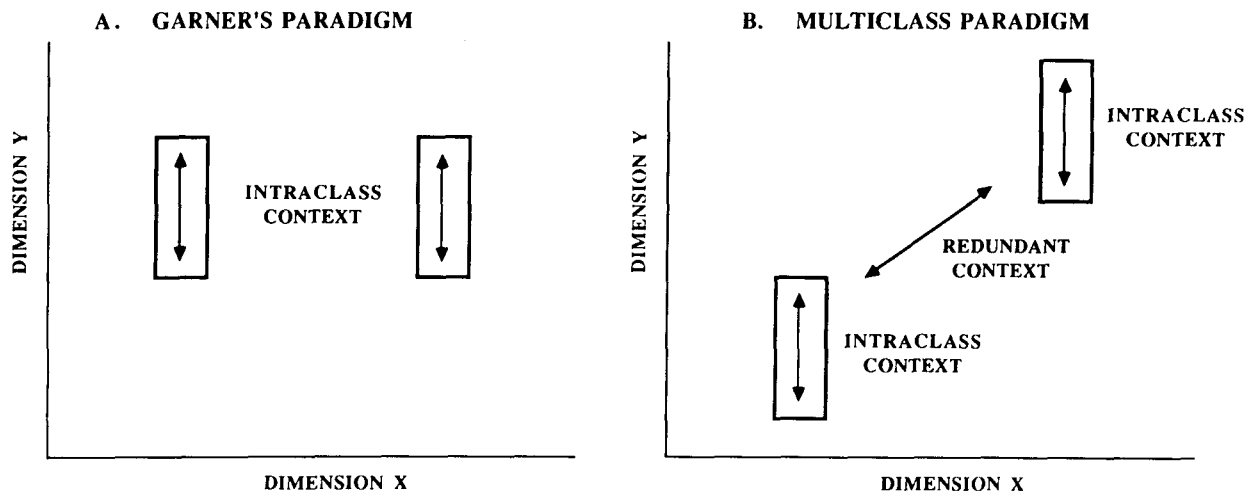


Figure 2. A schematic representation of Garner's filtering task (panel A), which manipulates only intraclass context, and a multiclass task (panel B), which combines intraclass context with redundant context.

different effects on performance. By comparing performance in dual-context tasks with measures of baseline performance, we are able to estimate the relative effects of these contexts on classification of attributes from a given dimension.

We assume that perceptual dimensions differ reliably with respect to how dual context affects them. In the present study, we found that, for any given member of a pair of auditory dimensions, one kind of context appears to dominate perception of attributes over the other kind. Thus, with the dimensions pitch and loudness, we found that, when one classifies attributes of pitch, the positive effects of redundant context appear to dominate performance, whereas, when one classifies attributes of loudness, the ill effects of intraclass context dominate. For simplicity, we here distinguish dimensions according to whether or not attributes appear to resist the effects of intraclass context. **HARD** dimensions are interacting dimensions that display such resistance. **SOFT** interacting dimensions are those that are greatly affected by intraclass variability; that is, they display substantial impairment from that context. To the best of our knowledge, we are the first to identify such differences between perceptual dimensions using a timed classification procedure.

### The Multiclass Tasks

Figure 3 presents the stimulus space used to create sets for the series of multiclass tasks. The stimulus dimensions here are frequency (pitch) and intensity (loudness). There are two kinds of baseline tasks, both borrowed from Garner's paradigm. The *standard-baseline* tasks are Garner's single-dimension tasks: Subjects discriminate

values on one dimension (e.g., pitch) at a constant level of the other dimension (e.g., Stimulus A<sub>1</sub> vs. Stimulus D). The *correlated-baseline* task is Garner's correlated-dimensions task; the stimulus set contains A<sub>1</sub> and B<sub>1</sub>.

We are most interested in how these baseline measures compare with performance in *one-way classification tasks*. In one-way tasks, each set contains four or six stimuli, and all stimuli in one category contain the same value on the criterial dimension. For example, in a one-way pitch task, Stimuli A<sub>1</sub> and A<sub>5</sub> may be assigned to one category and B<sub>1</sub> and B<sub>5</sub> to the other. Subjects are asked to distinguish high pitch (950 Hz: B<sub>1</sub> and B<sub>5</sub>) from low pitch (900 Hz: A<sub>1</sub> and A<sub>5</sub>). Within each category, stimuli vary on the irrelevant dimension (i.e., loudness: A<sub>1</sub> = 60 dB; A<sub>5</sub> = 50 dB), but this dimension is redundant with the relevant dimension. In our example, the high-pitch attribute is paired with one of the two loud attributes; the low-pitch attribute is paired with one of the two soft attributes.

By comparing performance in one-way tasks with that in standard-baseline and correlated-baseline tasks, we are able to determine which of the two types of context—redundant or intraclass—dominates perception of values on any particular criterial dimension. Slow performance relative to baseline implies that classification was disrupted significantly by intraclass variability, that is, that the dimension is **SOFT**. Fast performance relative to baseline implies that classification was not impaired particularly by intraclass context, but that it was instead enhanced due to redundancy, that is, that the dimension is **HARD**.

We also compare classification performance in one-way tasks with that in *two-way tasks*—that is, with performance in classifying stimuli whose dimensions are perfectly correlated linearly. In one of the two-way tasks, for example, Stimuli A<sub>1</sub> and A<sub>7</sub> are assigned to one category and Stimuli B<sub>1</sub> and B<sub>7</sub> to the other. Subjects are instructed to classify along both dimensions, because the two dimensions are equally valid predictors of category membership. The use of two-way tasks gives us an opportunity to compare (1) performance when values on the criterial dimension are equal to the number of categories (i.e., one-way tasks) with (2) performance when values on each dimension exceed the number of categories (i.e., two-way tasks). Some have suggested that subjects display a preference for categories distinguished by simple one-dimensional rules (e.g., Imai & Garner, 1965; Martin & Caramazza, 1980; Medin, Wattenmaker, & Hampson, 1987). We will examine whether such preferences translate into superior classification performance.

Finally, the multiclass paradigm permits interesting comparisons among RTs to specific stimuli. An important feature of the multiclass procedure is that in the correlated-baseline task and in each of the classification tasks, Stimulus A<sub>1</sub> is assigned to one response key and Stimulus B<sub>1</sub> to the other key. Thus we are able to examine speed to respond to these two *standard stimuli* (A<sub>1</sub> and B<sub>1</sub>) under various task conditions. Are standard stimuli embedded in a classification task processed as quickly as these same stimuli alone in a discrimination

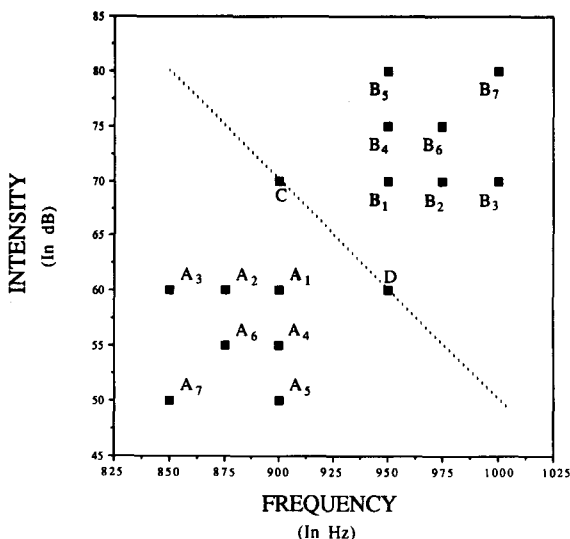


Figure 3. The stimulus space used to create sets for the series of multiclass tasks in experiments testing the dimensions pitch (frequency) and loudness (intensity) (Experiments 1A, 1B, and 2). Note—A reference boundary cuts through values for Stimulus C and Stimulus D, thus separating Category A from Category B (see text for explanation).

task? As we shall see, the answer to this question appears to depend on whether the classification dimension is **HARD** or **SOFT**.

We can also compare the RTs to the two standard stimuli with the RTs to individual stimuli within each category (e.g.,  $A_4$  and  $B_4$ ). This enables us to examine how classification performance changes as a function of the distance at which stimuli lie from some *reference boundary*—that is, from an imaginary line cutting through the points occupied by Stimulus C and Stimulus D, separating Category A from Category B (see Figure 3).<sup>1</sup>

In sum, then, multiclass permits us to explore effects of dual context on classification through an elaborate set of comparisons, both among tasks and among stimuli. We will find that these comparisons provide a rich new source of data about how perceivers process auditory dimensions.

## EXPERIMENTS 1A AND 1B

Melara and Marks (in press) showed that pitch and loudness interact integrally: Varying these dimensions orthogonally yields Garner interference, and correlating these dimensions produces redundancy gain (see also Grau & Kemler Nelson, 1988). Melara and Marks matched baseline discriminabilities of the two dimensions and found that effects of intraclass context alone (Garner interference) and redundant context alone (redundancy gain) were equivalent; that is, they were equal for the two dimensions. In other words, we showed that in single context, pitch and loudness behave equivalently. In Experiments 1A and 1B of the present study, we examined how pitch and loudness are affected by dual context. To anticipate, we found that pitch is a **HARD** dimension and loudness is a **SOFT** dimension. That is, pitch resists intraclass context (i.e., it is classified quickly relative to baseline), but loudness assimilates into intraclass context (i.e., it is classified slowly relative to baseline). So whereas pitch and loudness show comparable effects in Garner's paradigm, in the dual-context multiclass paradigm introduced here, variations on these same dimensions produce opposite effects.

In these experiments, we also compare speeds to respond to stimuli at different relative positions within a category (e.g., Stimuli  $A_1$  and  $B_1$  vs. Stimuli  $A_5$  and  $B_5$ ). As a working model, we use a simple distance to boundary: Stimuli that lie relatively close to our reference boundary (e.g.,  $A_1$ ) should be classified more slowly than stimuli that are distant from the boundary (e.g.,  $A_5$ ). Our working model is logically consistent with other geometric models of similarity/categorization (e.g., Krumhansl, 1978; Kruskal, 1964; Shepard, 1962a, 1962b; Torgerson, 1958). As this research unfolds, however, we will see that when certain dimensions are paired, this simple distance model is clearly inadequate.

Experiments 1A and 1B differed only in the instructions subjects received. The subjects in Experiment 1A were explicitly told the dimension or dimensions along

which classifications should be made. The subjects in Experiment 1B practiced each task before performing experimental trials, but they were never given explicit rules for classifying. We examine differences in performance resulting from the use of these different instructions.

## Method

**Subjects.** Twenty-five subjects (14 men and 11 women), mostly students from the Yale community, were paid \$5 per hour to participate. Fourteen subjects participated in Experiment 1A, 11 in Experiment 1B.

**Stimuli and Apparatus.** A total of 16 sounds were generated by the SID chip of a Commodore 64 microcomputer, directed to a Realistic SA-150 stereo amplifier, and presented over Realistic Nova-40 headphones. The frequency and intensity values for these sounds appear in Figure 3. The frequency varied between 850 and 1000 Hz, a relatively flat range on Fletcher-Munson equal loudness contours (Fletcher & Munson, 1933). Within each category, adjacent stimuli differed by 25 Hz; between categories, adjacent stimuli differed by 50 Hz. Intensity ranged between 50 and 80 dB. Within each category, adjacent stimuli differed by 5 dB, and between categories, they differed by 10 dB. Sound intensity was calibrated on the A-scale of a GenRad 1565-B sound-level meter. We selected these values of frequency and intensity because our earlier work suggested that in this range the dimensions pitch and loudness were equally discriminable. All stimuli were rectangular waveform tones having a duty cycle of .4620 (where .50 represents a square wave).

From these 16 sounds, we created either 14 (Experiment 1A) or 11 (Experiment 1B) experimental tasks of 48 trials each. Five were baseline discrimination tasks (4 standard baseline, 1 or 2 correlated baseline), and 6 were classification tasks (2 one-way tasks involving pitch, 2 one-way tasks involving loudness, and 2 or 4 two-way tasks). Equal numbers of the two stimuli (discrimination) or four stimuli (classification) were included in each task.

**Standard baseline.** Four stimulus sets each contained two stimuli that differed along one dimension but were held constant at some value on the other dimension. In two sets, pitch varied,  $A_1$  versus  $D$  and  $B_1$  versus  $C$ , and in two sets loudness varied,  $A_1$  versus  $C$  and  $B_1$  versus  $D$  (see Figure 3).

**Correlated baseline.** Subjects discriminated Stimulus  $A_1$  from Stimulus  $B_1$ . This task therefore involved a positive correlation between pitch and loudness. In addition, these two stimuli were included in all stimulus sets involving classification.

**One-way pitch.** All stimulus sets used in classification tasks contained four stimuli, two assigned to one category and two to the other category. Two one-way pitch tasks were tested, one involving a relatively large physical distance between stimuli within each category (large-width task), the other involving a smaller category width (small-width task). In small-width tasks, Stimuli  $A_1$  and  $A_4$  were assigned to one category and Stimuli  $B_1$  and  $B_4$  were assigned to the other. In large-width tasks, the A category contained Stimuli  $A_1$  and  $A_5$ , whereas the B category contained Stimuli  $B_1$  and  $B_5$ . In both stimulus sets, all stimuli in the A category were 900 Hz and varied in intensity from 50 to 60 dB; all B stimuli were 950 Hz and varied in intensity from 70 to 80 dB.

**One-way loudness.** Two stimulus sets, one involving small width and one involving large width, were created such that all stimuli in the A category were 60 dB, their frequency varying between 850 and 900 Hz, and all stimuli in the B category were 70 dB, their frequency varying between 950 and 1000 Hz. In the small-width task, subjects classified  $A_1$  and  $A_2$  versus  $B_1$  and  $B_2$ . In the large-width task, subjects classified  $A_1$  and  $A_3$  versus  $B_1$  and  $B_3$ .

**Two-way.** These stimulus sets were created by correlating pitch and loudness linearly. Thus, within each category, both dimensions

were varied. In the small-width task, subjects classified  $A_1$  and  $A_6$  versus  $B_1$  and  $B_6$ . In the large-width task, subjects classified  $A_1$  and  $A_7$  versus  $B_1$  and  $B_7$ .

**Procedure.** All subjects were tested individually in a darkened, sound-attenuating booth. The subjects in Experiment 1A were told to attend to one of the dimensions when performing each task. Half of these subjects first performed a set of tasks in which pitch was criterial, and half first performed tasks in which loudness was criterial. For these subjects, the correlated-baseline task and the two-way tasks occurred under both conditions—that is, when attention was to be focused on pitch and when it was to be focused on loudness. The set of tasks for the pitch condition also included the two standard-baseline tasks involving pitch and the two one-way pitch tasks. Similarly, the set of tasks administered when loudness was criterial included the baseline tasks involving loudness and the two one-way loudness tasks. Thus, each subject in Experiment 1A performed 14 different tasks, 7 in the pitch condition and 7 in the loudness condition. The subjects performed two consecutive runs of each of these 14 tasks. The order of tasks within each set of 7 was specified by two  $7 \times 7$  Latin squares. Each task began with 6–12 trials of practice. No practice data were analyzed.

The subjects in Experiment 1B were given no explicit rules for classifying. They were told simply that stimuli had been assigned to different response keys and that they should classify accordingly. These subjects performed 11 different tasks (the correlated-baseline task and the 2 two-way tasks were tested just once), their order specified by an  $11 \times 11$  Latin square. Again, each task was performed in two consecutive runs and was preceded by practice trials.

In both experiments, trials in each task were presented randomly (experimental or practice). The subject initiated each trial by pressing the space bar of the computer keyboard. The stimulus followed 1 sec later. Decisions were made by pressing either a right-hand or a left-hand key on the keyboard. Key assignment was counterbalanced across subjects. Time elapsed between stimulus onset and keypress was measured in milliseconds, using a software timer (Price, 1979). Response terminated the stimulus. The subjects were instructed to make their decisions as quickly as possible without error. A message displayed on the computer's monitor informed the subjects of incorrect responses. In addition, subjects responding slower than 800 msec were encouraged (through a screen message) to respond more quickly. Following each task, the subjects were informed of their mean RT and error rate. The subjects were encouraged to rest between tasks whenever they wished. Experiment 1A lasted approximately 2.5 h. Experiment 1B lasted approximately 2 h.

## Results

**Overall performance.** Mean RTs and error rates in each task appear in Table 1. The overall error rate was .05 in Experiment 1A and .04 in Experiment 1B. Errors averaged across conditions correlated .40 with RTs in Experiment 1A and .87 with RTs in Experiment 1B. Thus, it seems likely that speed and accuracy did not trade off with each other in either of these experiments. Errors were not analyzed further. Loudness was discriminated slightly faster than pitch at baseline in Experiment 1A [14 msec: 289 vs. 303 msec;  $F(1,13) = 4.26$ ,  $p = .06$ ], but these baseline discriminabilities did not differ reliably in Experiment 1B [312 vs. 319 msec;  $F(1,10) = .89$ , n.s.]. For our purposes in these experiments, noteworthy is the finding that pitch was not more discriminable than loudness in either experiment.

**Analyses involving tasks.** A redundancy gain was found in both experiments; that is, performance in the correlated-baseline task was significantly faster than that in the standard-baseline tasks [Experiment 1A (12 msec: 284 vs. 296 msec),  $F(1,13) = 7.07$ ,  $p = .02$ ; Experiment 1B (15 msec: 301 vs. 316 msec),  $F(1,10) = 17.53$ ,  $p = .002$ ]. (In Experiment 1A, redundancy gain was present for both pitch and loudness correlated baselines when compared separately with their corresponding standard baselines.) This replicates our previous finding with these dimensions (Melara & Marks, in press; see also Grau & Kemler Nelson, 1988).

The most important analysis here concerns relationships between discrimination (standard baseline) and classification (one-way) tasks in pitch and loudness judgments. To examine this, we tested the interaction between tasks (standard baseline, one-way) and dimensions (pitch, loudness), using analysis of variance (ANOVA). In both Experiments 1A and 1B, this interaction was significant [ $F(1,13) = 24.32$ ,  $p < .001$ , and  $F(1,10) = 12.06$ ,  $p < .01$ , respectively]. Inspection of Table 1 reveals the nature of these interactions. In both experiments, one-way

Table 1  
Mean Reaction Times (in milliseconds) to Perform One-Way Pitch,  
One-Way Loudness, and Two-Way Tasks (Experiments 1A and 1B)

Task	One-Way						Overall	
	Pitch		Loudness		Two-Way			
	M	Proportion of Error	M	Proportion of Error	M	Proportion of Error		
Experiment 1A								
Standard baseline	303	.09	289	.07			296	.08
Correlated baseline	288	.04	280	.04			284	.04
Small width	288	.05	301	.07	316	.05	301	.05
Large width	283	.06	316	.07	321	.08	306	.07
Experiment 1B								
Standard baseline	319	.04	312	.03			316	.04
Correlated baseline							301	.03
Small width	301	.03	356	.04	334	.04	330	.04
Large width	334	.03	360	.06	361	.06	352	.05

loudness classifications (1A = 309 msec, 1B = 358 msec) were slower than baseline loudness identifications [1A = 289 msec, 1B = 312 msec;  $F(1,13) = 13.58, p < .01$ , and  $F(1,10) = 18.23, p < .01$ , respectively]. With pitch judgments, however, classification was either marginally faster than discrimination [1A: 286 vs. 303 msec,  $F(1,13) = 3.64, p = .08$ ], or else the two kinds of tasks were equivalent [1B: 317 vs. 319 msec,  $F(1,10) = .05$ , n.s.]. In both experiments, one-way pitch classifications were performed significantly faster than one-way loudness classifications [ $F(1,13) = 5.49, p = .04$ , and  $F(1,10) = 12.06, p < .01$ , respectively].

Performance in two-way tasks (1A = 319 msec, 1B = 348 msec) was slower than performance in standard-baseline tasks in both experiments [ $F(1,13) = 7.29, p = .02$ , and  $F(1,10) = 9.91, p = .01$ , respectively]. Only in Experiment 1A, however, did overall one-way classifications (1A = 297 msec, 1B = 338 msec) differ significantly from two-way classifications [ $F(1,13) = 18.09, p < .01$ , and  $F(1,10) = 2.71$ , n.s.]. Still, in both experiments, one-way pitch classifications were faster than either one-way loudness or two-way classifications, as tested by Newman-Keuls comparisons [ $F(1,13) = 8.71, p < .01$ , and  $F(1,10) = 9.60, p < .01$ , respectively].

The two experiments differed in terms of the effects of category width on classification. In Experiment 1A, tasks involving large category width (306 msec) were performed virtually as quickly as those involving small width [301 msec;  $F(1,13) = 1.46$ , n.s.]. In Experiment 1B, on the other hand, small-width tasks (330 msec) were classified 22 msec faster on the average than were large-width tasks [ $F(1,10) = 21.45, p < .01$ ].

**Analyses involving stimuli.** A series of ANOVAs were used to examine differences in speed to respond to the two standard stimuli (i.e., Stimuli A<sub>1</sub> and B<sub>1</sub> in Figure 2). The mean RTs to classify these stimuli in the various tasks appear in Table 2. In Experiment 1A, RT to classify the standard stimuli in the one-way pitch tasks (293 msec) was equivalent to RT to discriminate these stimuli by pitch in the correlated-baseline task [288 msec;  $F(1,13) = .86$ , n.s.]. But in Experiment 1B, standard stimuli were classified 30 msec slower in one-way pitch tasks (331 msec) than in the correlated-baseline task [301 msec;  $F(1,10) = 28.5, p < .001$ ]. Subjects were slower to classify the standard stimuli in one-way loudness tasks than to discriminate them in correlated-baseline tasks in both experi-

ments [ $F(1,13) = 50.98, p < .001$ , and  $F(1,10) = 34.61, p < .001$ , respectively]. Similarly, in both experiments, standard stimuli were classified slower in two-way tasks (1A = 331 msec, 1B = 359 msec) than discriminated in the correlated-baseline task [ $F(1,13) = 39.47, p < .001$ , and  $F(1,10) = 52.59, p < .001$ , respectively].

The subjects in both experiments responded to standard stimuli faster in one-way pitch tasks (1A = 293 msec, 1B = 331 msec) than in either one-way loudness tasks (1A = 314 msec, 1B = 365 msec) or in two-way tasks (1A = 331 msec, 1B = 359 msec), as tested by ANOVA and Newman-Keuls [ $F(2,26) = 9.11, p < .01$ , and  $F(2,20) = 5.23, p = .02$ , respectively].

Finally, in all classification tasks and in both experiments, the subjects responded to standard stimuli (1A = 313 msec, 1B = 352 msec) more slowly than to the other stimuli in the category [1A = 298 msec, 1B = 330 msec;  $F(1,13) = 73.8, p < .001$ , and  $F(1,10) = 44.84, p < .001$ , respectively; see Table 2].

## Discussion

Experiments 1A and 1B provided a rich set of findings. First and foremost, we found that pitch and loudness differ in how they were affected by dual context. Classification of loudness was impaired by the intraclass variation produced by changes in pitch, but classification of pitch resisted interference from intraclass variation produced by changes in loudness. Thus, pitch may be thought of as a **HARD** integral dimension because it resists intraclass context, whereas loudness may be thought of as a **SOFT** integral dimension, because it is readily assimilated into intraclass context.

The importance of these findings obtained in dual context is seen clearly when one compares them with findings obtained in single-context classification of the same dimensions (e.g., see Melara & Marks, in press). Effects obtained previously in Garner's paradigm were almost perfectly symmetric: Pitch helped (redundancy gain) or hindered (Garner interference) performance when loudness was relevant, as much as loudness helped or hindered performance when pitch was relevant. Thus, when redundant context or intraclass context was present alone, as they are in Garner's procedure, these dimensions were indistinguishable. But in the dual-context multiclass procedure, effects of the irrelevant dimension on performance were not symmetric. Subjects were significantly better at focusing on values of pitch when loudness varied than at focusing on loudness when pitch varied. Thus, when redundant context and intraclass context are present simultaneously, we discover that all integral dimensions are not equal. Pitch and loudness interact symmetrically in single context (see Garner, 1976). But in dual context, pitch and loudness interact asymmetrically. To the best of our knowledge, these findings are completely new.

**Effects of strategy.** Several findings here suggested that classification under explicitly analytic instructions differed in important ways from classification when instructions were inexplicit. In Experiment 1A, when instructions were explicit, subjects classified pitches faster than they

Table 2  
Mean Reaction Times (in milliseconds) to Stimuli in One-Way Pitch, One-Way Loudness, and Two-Way Tasks (Experiments 1A and 1B)

Stimuli	One-Way		Two-Way	Overall
	Pitch	Loudness		
Experiment 1A				
Standard	293	314	331	313
Other	279	303	312	298
Experiment 1B				
Standard	331	365	359	352
Other	304	352	334	330

discriminated them; performed one-way tasks faster than two-way tasks; showed no effect of within-category distance on classification; and classified standard stimuli equally fast in the one-way pitch tasks and in the correlated-baseline task. In Experiment 1B, on the other hand, classification and discrimination of pitch were performed equivalently, as were one-way tasks relative to two-way tasks; classification was slower when category width was large versus small; and RT to standard stimuli was faster in correlated-baseline discrimination than in one-way pitch classification.

The difference between the two types of instructions appears largely to be an advantage in performance reaped by the analytic instructions, rather than due to any interaction between instructions and conditions (e.g., nonanalytic instructions being better under some conditions, analytic better under others). This advantage was especially pronounced for the HARD dimension, pitch, presumably because attentive focus enabled subjects to avoid deleterious effects of category width when classifying according to that dimension. Still, no matter which dimension was classified, larger category width impaired performance (relative to small width) only when instructions were nonanalytic. Similarly, both one-way tasks (relative to two-way tasks) benefited from analytic over nonanalytic instructions. At least in the multiclass procedure, then, nonanalytic instructions seem to foster a mode of processing that is less efficient than analytic instructions.

These results fit well with our general approach for investigating how dimensions interact. Two aspects of the approach are relevant in this regard. First, in our view, perception of the stimulus whole derives from the effect of context on the perception of attributes (see Melara & Marks, *in press*). So, when one studies stimulus perception, it seems sensible to stress how subjects classify attributes when told explicitly to do so. Second, we believe that the best way to understand how subjects perceive multidimensional stimuli is to study performance under conditions optimal for subjects. The results of these experiments suggest that analytic instructions lead to benefits but no costs. Consequently, in the experiments that follow, multiclass tasks are performed under explicit instructions.

**Classifying standard stimuli.** Analyses involving standard stimuli revealed two primary determinants of classification performance. First, these stimuli were classified faster when the task involved a one-way classification of the HARD dimension (pitch) than when the task was either one-way involving the SOFT dimension (loudness) or two-way. These results are consistent with the between-task analyses already discussed. The results suggest that the same stimulus may be perceived differently, depending on (1) whether the criterial dimension is HARD or SOFT (i.e., whether the dominant context is redundant context or intraclass context) and (2) where subjects' attention is directed. For instance, subjects benefited from redundant context when attention was directed to the HARD dimension and values on that dimension were constant within

category. But subjects suffered from intraclass context when attention was directed to the SOFT dimension despite constant values on that dimension within category.

The analyses of standard stimuli point, second, to the relevance of between-category distance in accounting for classification performance. Standard stimuli were closer physically to our reference boundary (see Figure 3) than were other stimuli within a category. Consistent with this proximity, standard stimuli were classified more slowly than other stimuli. Thus it appears that, at least with these dimensions, the farther a stimulus is from our reference boundary (i.e., the farther from the other category), the more efficient is classification.

## EXPERIMENT 2

In Experiment 2, we had two goals. First, we wished again to dissociate HARD (pitch) from SOFT (loudness) dimensions, using explicit instructions. In particular, we were interested in replicating with different subjects the diverse pattern of findings obtained in Experiment 1A.

Second, we wished to examine performance in a set of classification conditions where each category contained three stimuli. Do such three-stimulus categories lead inevitably to worse performance than two-stimulus categories or than discriminations (because of the greater trial-to-trial variability), or does overall performance in such tasks depend more on the HARDNESS of the attended dimension? Are RTs to the three stimuli within a category a monotonic function of the distance at which those stimuli lie from the reference boundary? To answer these questions, the subjects in Experiment 2 performed one-way pitch tasks, one-way loudness tasks, and two-way tasks for stimulus sets containing six stimuli, three assigned to each category.

### Method

**Subjects.** Twenty subjects (12 men and 8 women) were paid \$5 per hour to participate.

**Stimuli, Apparatus, and Procedure.** The apparatus and stimuli were identical to those in the previous experiments. The subjects performed two consecutive runs of 10 experimental tasks, their order specified by two  $10 \times 10$  Latin squares. These tasks involved a set of three one-way pitch tasks, a set of three one-way loudness tasks, and a set of three two-way tasks. Within each set of three tasks, one task contained two stimuli in a category and involved small width, one involved two-stimulus categories and large width, and one involved stimuli mapped onto each of two response categories. In the three-stimulus classifications, a category comprised the three stimuli grouped together in Figure 3 in a row (one-way loudness), in a column (one-way pitch), or on a diagonal (two-way). So, for example, in the one-way loudness task, Category A was the row containing Stimuli A<sub>1</sub>, A<sub>2</sub>, and A<sub>3</sub>, whereas Category B was the row containing Stimuli B<sub>1</sub>, B<sub>2</sub>, and B<sub>3</sub>. We refer to stimuli positioned adjacent to the standard stimuli in three-stimulus categories (e.g., A<sub>2</sub> and B<sub>2</sub>) as *adjacent* stimuli, and stimuli positioned farthest from the standard stimuli/reference boundary (e.g., A<sub>3</sub> and B<sub>3</sub>) as *distant* stimuli.

In addition to the nine classification tasks, the subjects performed the correlated-baseline task. No standard-baseline tasks were used in this experiment. Prior to performing each experimental task, the

**Table 3**  
**Mean Reaction Times (in milliseconds) to Perform One-Way Pitch,**  
**One-Way Loudness, and Two-Way Tasks (Experiment 2)**

Category (Task)	One-Way		One-Way		Two-Way		Overall	
	Pitch		Loudness		Two-Way		Overall	
	<i>M</i>	Proportion of Error	<i>M</i>	Proportion of Error	<i>M</i>	Proportion of Error	<i>M</i>	Proportion of Error
Correlated baseline							287	.05
Two-stimulus (small width)	280	.04	309	.05	304	.06	299	.05
Two-stimulus (large width)	285	.05	327	.08	323	.08	312	.07
Three-stimulus	278	.05	318	.07	303	.06	299	.06
Mean	281	.05	318	.06	310	.07		

subjects read a set of instructions directing them explicitly to classify either according to the pitch dimension (for one-way pitch tasks) or the loudness dimension (for one-way loudness tasks). For correlated-baseline and two-way tasks, the subjects were told to classify by either pitch or loudness; thus, these tasks were tested in the same number of trials as one-way tasks were. All other aspects of the procedure duplicated those of the previous experiments.

### Results

**Overall performance.** RTs and error rates in each task appear in Table 3. Across conditions, RTs and errors correlated .86. The overall error rate was .05. Errors were not analyzed further.

**Analyses involving tasks.** To distinguish HARD from SOFT in these analyses, performance in the correlated-baseline condition (287 msec) was used as the comparison. Even though this experiment did not include standard-baseline tasks, we assumed that performance in the correlated-baseline task reflected a gain from correlational redundancy.

Performance in the correlated-baseline task did not differ from that in the pooled set of (three) one-way pitch tasks [281 msec;  $F(1,19) = 1.61$ , n.s.]. When this set was broken down into the tasks involving two-stimulus categories (283 msec) and those involving three-stimulus categories (278 msec), we found that the former was equivalent to baseline [ $F(1,19) = .56$ , n.s.], but that the latter was statistically faster than baseline [ $F(1,19) = 4.63$ ,  $p = .05$ ].

Performance in the correlated-baseline task was significantly faster than that in the pooled set of (three) one-way loudness tasks [318 msec;  $F(1,19) = 25.43$ ,  $p < .001$ ]. This difference was identical for the two-stimulus (318 msec) and three-stimulus (318 msec) categories [ $F(1,19) = 24.62$ ,  $p < .001$ , and  $F(1,19) = 22.04$ ,  $p < .001$ , respectively].

The correlated-baseline task was also performed somewhat faster than the pooled two-way tasks [310 msec;  $F(1,19) = 12.77$ ,  $p < .01$ ]. This difference was greater, however, for two-stimulus categories [314 msec;  $F(1,19) = 16.56$ ,  $p = .001$ ] than for three-stimulus categories [303 msec;  $F(1,19) = 3.5$ ,  $p = .08$ ]. RTs in pooled one-way pitch, one-way loudness, and two-way tasks differed significantly [ $F(2,38) = 27.1$ ,  $p < .001$ ]. A Newman-Keuls post hoc analysis (.05 criterion level) revealed that attributes were classified faster in one-way pitch tasks than

in either one-way loudness tasks or two-way tasks, which did not differ from each other. Overall, attributes were classified slightly faster in one-way tasks (300 msec) than in two-way tasks [310 msec;  $F(1,19) = 6.75$ ,  $p = .02$ ]. Finally, in two-stimulus categories, stimuli were classified faster when category width was small (299 msec) than when it was large [312 msec;  $F(1,19) = 7.98$ ,  $p = .01$ ].

**Analyses involving stimuli.** For these analyses, RT to discriminate the standard stimuli in the correlated-baseline task (287 msec) was compared with RT to classify standard stimuli in various classification tasks. These RTs appear in Table 4. With one-way pitch tasks (291 msec), discrimination and classification of standard stimuli did not differ [ $F(1,19) = .84$ , n.s.]. With one-way loudness tasks (327 msec), on the other hand, the standard stimuli were classified 40 msec slower than they were discriminated [ $F(1,19) = 34.54$ ,  $p < .001$ ]. Similarly, with two-way tasks (328 msec), classification was significantly slower than discrimination [ $F(1,19) = 27.53$ ,  $p < .001$ ]. An ANOVA and subsequent Newman-Keuls analysis of RTs to standard stimuli in one-way pitch, one-way loudness, and two-way tasks revealed that in one-way pitch tasks standard stimuli were classified faster than they were in either of the other tasks, which did not differ from each other [ $F(2,38) = 21.02$ ,  $p < .001$ ].

The foregoing effects were all consistent for two-stimulus or three-stimulus categories examined alone. Furthermore, in two-stimulus categories, standard stimuli (315 msec) were always classified slower than the other stimulus in the category [294 msec;  $F(1,19) = 61.26$ ,  $p < .001$ ; see Table 4]. Similarly, in three-stimulus

**Table 4**  
**Mean Reaction Times (in milliseconds) to Stimuli in One-Way Pitch,**  
**One-Way Loudness, and Two-Way Tasks (Experiment 2)**

One-Way Loudness, and Two-Way Tests (Experiment 2)				
Stimuli	One-Way		Two-Way	Overall
	Pitch	Loudness		
Two-Stimulus Category				
Standard	291	327	328	315
Other	274	310	300	294
Three-Stimulus Category				
Standard	293	329	332	318
Adjacent	274	316	289	293
Distant	266	309	288	287



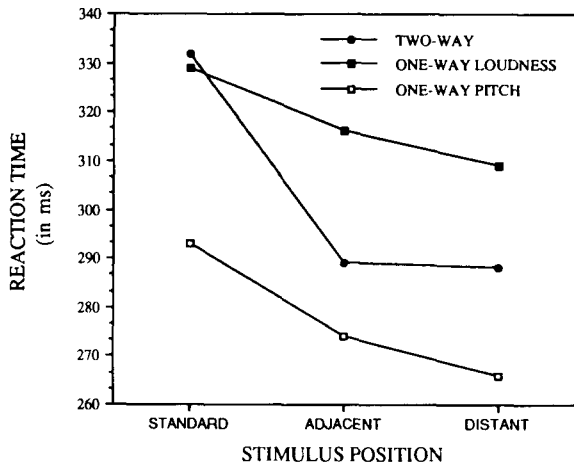


Figure 4. Mean reaction time (in milliseconds) to classify standard, adjacent, and distant stimuli in one-way pitch, one-way loudness, and two-way tasks (Experiment 2).

categories, standard stimuli (318 msec) were classified slower than adjacent stimuli (293 msec) or distant stimuli [287 msec;  $F(1,19) = 36.09, p < .001$ ]. Newman-Keuls analysis showed that adjacent and distant stimuli did not differ from each other (see Figure 4).

### Discussion

These results are largely in keeping with those from Experiment 1A. Pitch and loudness—which were matched previously in baseline discriminability and which were shown previously to yield equivalent amounts of Garner interference (Melara & Marks, *in press*)—lead to an opposite pattern of performance in the multiclass paradigm. In the present experiment, pitch classification was never worse than correlated-baseline discrimination under any condition. Moreover, RT to standard stimuli was relatively constant across pitch classification and baseline discrimination tasks. Loudness classification, however, was always worse than baseline under all conditions. Standard stimuli were responded to more slowly when embedded in a loudness classification task than when embedded in a baseline discrimination task.

We interpret these results as reflecting differences in how dual context affects interacting dimensions. We argue that attributes of loudness were veiled perceptually in dual context, and so classification performance suffered. Attributes of pitch, on the other hand, resisted the ill effects of intraclass context, and benefited from the redundancy between dimensions. We add, however, that even classification of the HARD dimension, pitch, probably suffered some cost from intraclass context, because no classification task was performed substantially better than any discrimination task, despite the fact that the additional stimuli in the classification tasks were located farther out in the stimulus space. Thus, perhaps, interacting dimensions, unlike separable dimensions, can never com-

pletely avoid interference from intraclass context, even when redundant context is present simultaneously.

The beneficial effects of redundant context for pitch, but not for loudness, were particularly apparent when categories contained three stimuli (see Figure 4). In these tasks, both intraclass variation and redundant variation were greater. For pitch, the redundant component dominated, leading to classification performance that surpassed even correlated-baseline discrimination. For loudness, classification was equivalent for two-stimulus and three-stimulus categories, suggesting that either (1) interference from intraclass variation had reached ceiling level or (2) the additional redundant variation did have a beneficial effect, but that effect was negated by the additional intraclass variation. In any case, results from the classification of three-stimulus categories are compelling in distinguishing effects of intraclass context from redundant context and, hence, HARD from SOFT interacting dimensions.

The classification of three-stimulus categories also provided insights into the cognitive structure of these categories. As seen in Figure 3, the speed to classify stimuli decreased monotonically as distance increased from the reference boundary. Noteworthy is the finding that the relationship between speed and distance is not strictly linear; rather, the functions are more nearly negatively accelerating, especially the one depicting two-way judgments. This finding suggests that the category structure may have reorganized the similarity space. Physically, the standard and adjacent stimuli are as different from each other as the adjacent and distant are from each other. Thus, if we assume that RT is an approximately linear function of degree of perceptual similarity (see, e.g., Melara, 1989b; Podgorny & Garner, 1979), the fact that adjacent and distant stimuli are classified more rapidly than standard stimuli implies that the adjacent and distant stimuli have become perceptually less different from each other than the standard and adjacent are from each other. In the case of pitch and loudness, then, the category boundary appears to act like a magnet, exerting disproportionately greater influence on stimuli nearest to it. But as we shall see, the magnet metaphor is not appropriate for all pairs of interacting dimensions.

### EXPERIMENT 3

Experiments 1A, 1B, and 2 consistently resulted in differences between pitch classification and loudness classification under conditions of dual context. Are these effects peculiar to the pairing of pitch and loudness? Or is the HARD-SOFT distinction evident when other interacting dimensions are paired? To answer these questions, we chose a third dimension of sound, which has been shown to interact with both pitch and loudness. This dimension is timbre, the qualitative character of complex tones. In Experiment 3, we paired timbre with loudness in the multiclass procedure. Melara and Marks (1989)

have recently demonstrated that timbre and loudness interact integrally in Garner's speeded classification procedure: Intra-class variation (single context) of one dimension produced substantial Garner interference when classifying the other dimension. Importantly, the absolute amount of Garner interference produced was the same for timbre classifications as for loudness classifications. Thus, as with pitch and loudness, timbre and loudness behave equivalently when tested in Garner's paradigm. As we will see next, however, timbre and loudness are affected very differently by the dual context of the multi-class paradigm.

In Experiment 3, we used what we will hereafter regard as the complete multiclass paradigm. Each subject performed 14 different tasks: 4 standard-baseline tasks, 1 correlated-baseline task, 3 one-way timbre tasks, 3 one-way loudness tasks, and 3 two-way tasks. Hence we were able to provide an in-depth analysis of task effects and stimulus effects that obtain for classifying these dimensions.

## Method

**Subjects.** Fourteen subjects (9 men and 5 women) were paid \$5 per hour of participation.

**Stimuli, Apparatus, and Procedure.** The apparatus were identical to that used in previous experiments. The stimulus space used to create the experimental tasks appears in Figure 5. Our timbre dimension was the perceptual correlate of duty cycle; that is, it was the ratio of pulse width to cycle length of a variable pulse (rectangular) waveform. At a given frequency, this ratio was held con-

stant. Duty cycle is expressed as a proportion, with .50 representing a square wave. Duty cycle varied by .0625 steps within each category and ranged between .0628 (twangy sounding) and .4379 (hollow sounding); the two categories were separated by a width of .125. The stimuli were separated by intensity differences of 3 dB within category and 6 dB between categories; intensity ranged between 58 and 76 dB. All sounds had fundamental frequencies of 950 Hz.

The 14 multiclass tasks were constructed exactly as in the previous experiments. The order of tasks was specified by a  $14 \times 14$  Latin square. As before, subjects performed each task in two consecutive 48-trial runs. All other aspects duplicated those of the previous experiments.

## Results

**Overall performance.** RTs and error rates for each of the 14 tasks appear in Table 5. The overall error rate was .08; speed correlated .98 with accuracy. Baseline loudness judgments (319 msec) were performed 30 msec faster than baseline timbre judgments [349 msec;  $F(1,13) = 38.7, p < .001$ ].

**Analyses involving tasks.** One-way timbre tasks were performed 13 msec faster than baseline timbre judgments [ $F(1,13) = 4.38, p < .06$ ]. When categories contained two stimuli, one-way timbre tasks (343 msec) did not differ from baseline [ $F(1,13) = .67, n.s.$ ]. Only when categories contained three stimuli were timbre judgments (324 msec) performed reliably faster than baseline [ $F(1,13) = 13.97, p < .01$ ].

One-way loudness tasks were performed 68 msec slower than baseline loudness tasks [ $F(1,13) = 82.54$ ,

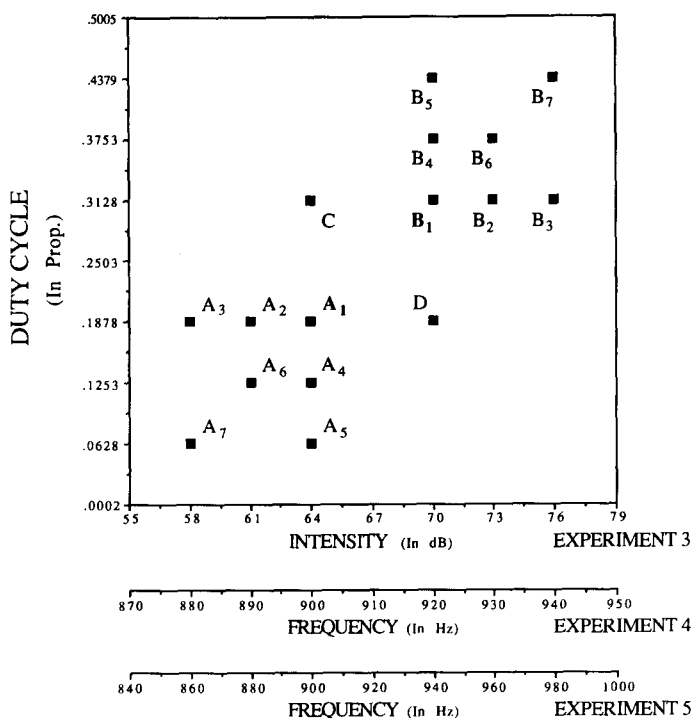


Figure 5. The stimulus space used to create multiclass tasks for testing the pair of dimensions timbre (duty cycle) and loudness (Experiment 3), and timbre and pitch (Experiments 4 and 5).

Table 5  
Mean Reaction Times (in milliseconds) to Perform One-Way Loudness,  
One-Way Timbre, and Two-Way Tasks (Experiment 3)

Category (Task)	One-Way Loudness		One-Way Timbre		Two-Way		Overall	
	Proportion of Error		Proportion of Error		Proportion of Error		Proportion of Error	
	M		M		M		M	
Standard baseline	319	.05	349	.07			334	.06
Correlated baseline							334	.07
Two-stimulus (small width)	380	.11	334	.06	359	.08	357	.08
Two-stimulus (large width)	404	.13	351	.08	373	.08	376	.10
Three-stimulus	377	.10	324	.06	373	.10	358	.09

$p < .001$ ]. This difference was significant both for two-stimulus categories [392 msec;  $F(1,13) = 86.25$ ,  $p < .001$ ] and for three-stimulus categories [377 msec;  $F(1,13) = 30.62$ ,  $p < .001$ ].

Two-way tasks (368 msec) were compared with the overall average of standard-baseline tasks (334 msec). This difference was significant [ $F(1,13) = 14.73$ ,  $p < .01$ ]. Two-way tasks were slower than baseline, whether two (366 msec) or three (373 msec) stimuli were present in a category [ $F(1,13) = 9.39$ ,  $p < .01$ , and  $F(1,13) = 16.23$ ,  $p < .01$ , respectively]. Two-way tasks did not differ from the overall RT to one-way tasks [362 msec;  $F(1,13) = .90$ , n.s.]. ANOVA and a Newman-Keuls post hoc test revealed that this was because one-way timbre tasks were performed significantly faster than two-way tasks, but one-way loudness tasks were performed significantly slower than two-way tasks [ $F(1,13) = 19.2$ ,  $p < .001$ ].

In tasks involving two-stimulus categories, small-width categories (357 msec) led to significantly faster responding than did large-width categories [376 msec;  $F(1,13) = 10.61$ ,  $p < .01$ ].

**Analyses involving stimuli.** Speed to classify the standard stimuli in the correlated-baseline task (334 msec) was used as the basis for all comparisons. RTs to classify individual stimuli in each classification task appear in Table 6. In one-way timbre tasks (349 msec), standard stimuli were classified 15 msec slower than they were discriminated, a statistically unreliable difference [ $F(1,13) = 3.28$ ,  $p = .09$ ]. In one-way loudness tasks (395 msec), on the other hand, standard stimuli were classified an appreciable 61 msec slower than they were discriminated [ $F(1,13) = 37.52$ ,  $p < .001$ ]. Standard stimuli were also

classified slowly in two-way tasks [389 msec;  $F(1,13) = 18.67$ ,  $p = .001$ ]. ANOVA and a subsequent Newman-Keuls test showed that RT to standard stimuli was faster in one-way timbre tasks than in either one-way loudness tasks or two-way tasks, which did not differ from each other [ $F(1,13) = 17.26$ ,  $p < .001$ ].

Among classification tasks with two-stimulus categories, standard stimuli (378 msec) were classified significantly slower than the other (more distant) stimuli in the category [355 msec;  $F(1,13) = 53.12$ ,  $p < .001$ ]. Among tasks with three-stimulus categories, standard stimuli (375 msec) were classified significantly slower than either stimuli adjacent to them in the category (354 msec) or stimuli distant from them in the category [345 msec;  $F(1,13) = 33.55$ ,  $p < .001$ ]. A Newman-Keuls test showed, additionally, that RTs to adjacent and distant stimuli differed reliably from each other. All of the above effects were found for one-way timbre, one-way loudness, and two-way tasks tested separately.

## Discussion

The pattern obtained here with timbre-loudness is strikingly similar to that obtained in Experiments 1A, 1B, and 2 with pitch-loudness. Baseline discriminability of loudness was vastly superior to that of timbre. Yet, in dual-context classification, loudness judgments were inferior to baseline performance, whereas timbre judgments were at or superior to baseline levels. These results suggest clearly that subjects can focus on the timbre dimension in the face of (redundant) intraclass variability, but that they have difficulty focusing on loudness. Thus, as in the case of pitch and loudness, we conclude that loudness is a SOFT dimension but that timbre is a relatively HARD dimension. The distinction between HARD and SOFT dimensions thus generalizes to a second pair of interacting dimensions, indicating that these findings are not idiosyncratic to pitch and loudness.

Other results also replicate those of previous experiments. We found no difference between one-way and two-way tasks, for example, but it was clear that this null finding was based in the HARDNESS of the dimension used in one-way classification. A one-way rule is more efficient than a two-way rule if the one-way dimension is HARD, but one-way is less efficient than two-way if the one-way dimension is SOFT. Thus, in the multiclass paradigm at least, the HARDNESS of the dimension is a better predic-

Table 6  
Mean Reaction Times (in milliseconds) to Stimuli in One-Way Loudness, One-Way Timbre, and Two-Way Tasks (Experiment 3)

Stimuli	One-Way		Two-Way	Overall
	Loudness	Timbre		
Two-Stimulus Category				
Standard	395	349	389	378
Other	379	324	362	355
Three-Stimulus Category				
Standard	391	335	399	375
Adjacent	367	323	372	354
Distant	373	313	349	345

**Table 7**  
**Mean Reaction Times (in milliseconds) to Perform One-Way Pitch,**  
**One-Way Timbre, and Two-Way Tasks (Experiment 4)**

Category (Task)	One-Way				Two-Way		Overall	
	Pitch		Timbre		Proportion of Error		Proportion of Error	
	<i>M</i>	Proportion of Error	<i>M</i>	Proportion of Error	<i>M</i>	Proportion of Error	<i>M</i>	Proportion of Error
Standard baseline	324	.08	322	.07			323	.08
Correlated baseline							291	.05
Two-stimulus (small width)	305	.06	283	.03	296	.05	294	.05
Two-stimulus (large width)	317	.07	296	.04	317	.06	310	.06
Three-stimulus	314	.06	300	.04	304	.05	306	.05

tor of classification performance than is the type of classificatory rule used.

Finally, we again found a set of effects based on distances between and within categories. Classification performance was better when the overall distance separating stimuli within a category was small versus large. And, as in previous experiments, stimuli closest to the reference boundary (standard stimuli) were classified slowest, no matter which dimension or rule was relevant to classification. As in Experiment 2, the difference in RT between the standard stimuli and adjacent stimuli was much greater than the difference between adjacent stimuli and distant stimuli (cf. Figure 4). Given the consistency across experiments, we regard the pattern of findings we have obtained here as the signature of perceptual interaction between a HARD and a SOFT dimension.

#### EXPERIMENT 4

In all of the experiments presented thus far, we have looked at the nature of interaction between a HARD dimension and a SOFT dimension. In the remaining experiments, we turn our attention to processes underlying interaction between two HARD dimensions, namely, pitch and timbre. As with timbre and loudness, Melara and Marks (1989) showed that, when paired, both timbre and pitch display substantial classification interference in the single context of intraclass variation. In Experiment 4, we use the complete multiclass procedure to examine effects of dual context on the classification of sounds varying in pitch and timbre.

#### Method

**Subjects.** Fourteen subjects (5 men and 9 women) were paid \$5 per hour to participate.

**Stimuli, Apparatus, and Procedure.** The apparatus and procedure were identical to those in Experiment 3. The stimulus space used to create the fourteen multiclass tasks appears in Figure 5. The values of duty cycle were the same as in Experiment 3. Frequency varied in 10-Hz steps within category, at least 20 Hz between categories; frequency ranged between 880 Hz and 940 Hz. Intensity was held constant at 70 dB.

#### Results

**Overall performance.** RTs and error rates appear in Table 7. The overall error rate was .06. Speed and accuracy correlated .90. Baseline discriminabilities of pitch (324 msec) and timbre (322 msec) matched well.

**Analyses involving tasks.** One-way pitch tasks (312 msec) were performed 12 msec faster than baseline, although this difference failed to reach significance [ $F(1,13) = 3.69, p = .08$ ]. Performance in classifying pitch in both two-stimulus categories (311 msec) and three-stimulus categories (314 msec) was numerically, though not statistically, faster than baseline [ $F(1,13) = 3.24, p = .10$ , and  $F(1,13) = 2.70, p = .12$ , respectively].

One-way timbre tasks (293 msec) were performed 29 msec faster than baseline timbre tasks [ $F(1,13) = 29.62, p < .001$ ]. Both two-stimulus categories (289 msec) and three-stimulus categories (300 msec) showed this gain [ $F(1,13) = 33.88, p < .001$ , and  $F(1,13) = 9.32, p < .01$ , respectively].

Performance in two-way tasks (306 msec) was faster than overall baseline performance [323 msec;  $F(1,13) = 11.26, p < .01$ ]. Again, both two-stimulus categories (307 msec) and three-stimulus categories (304 msec) showed the gain [ $F(1,13) = 6.69, p = .02$ , and  $F(1,13) = 10.72, p < .01$ , respectively]. Performance in two-way tasks did not differ from overall performance in one-way tasks [303 msec;  $F(1,13) = 1.36, n.s.$ ]. As in previous experiments, this was due to the fact that one of the one-way tasks, pitch, was performed slower than two-way tasks, and the other, timbre, was performed faster, although only the latter difference was reliable by Newman-Keuls [ $F(2,26) = 10.48, p < .001$ ].

Across all tasks involving two-stimulus categories, small-width tasks (294 msec) were performed significantly faster than large-width tasks [310 msec;  $F(1,13) = 10.64, p < .01$ ].

**Analyses involving stimuli.** RTs to classify individual stimuli in each of the three types of classification tasks—one-way pitch, one-way timbre, two-way—appear in Table 8. In one-way pitch tasks (312 msec), standard stimuli were classified 21 msec slower than they were discriminated in the correlated-baseline task [291 msec;  $F(1,13) = 8.11, p = .01$ ]. On the other hand, in one-way timbre tasks, standard stimuli were classified (290 msec) as fast as they were discriminated [ $F(1,13) = .03, n.s.$ ]. Standard stimuli were classified slowly in two-way tasks (311 msec) relative to correlated-baseline performance [ $F(1,13) = 9.93, p < .01$ ]. Standard stimuli were classified significantly faster in one-way timbre tasks than they were in either one-way pitch or two-way tasks, which did not differ from each other [ $F(2,26) = 9.68, p = .01$ ].

**Table 8**  
**Mean Reaction Times (in milliseconds) to Stimuli in One-Way Pitch, One-Way Timbre, and Two-Way Tasks (Experiment 4)**

Stimuli	One-Way		Two-Way	Overall
	Pitch	Timbre		
Two-Stimulus Category				
Standard	312	290	311	302
Other	313	298	295	302
Three-Stimulus Category				
Standard	313	296	315	308
Adjacent	320	306	302	309
Distant	310	300	296	302

In two-stimulus categories, there was no difference in speed to classify standard stimuli (302 msec) versus other stimuli in the category (302 msec). In three-stimulus categories, the most distant stimuli in the category (302 msec) were classified slightly faster than either the standard stimuli (308 msec) or the adjacent stimuli [309 msec;  $F(2,26) = 4.19, p = .03$ ]. However, stimulus position (standard, adjacent, distant) interacted with type of classification [one-way pitch, one-way timbre, two way;  $F(4,52) = 5.0, p < .01$ ]. Only in two-way tasks did performance improve with increasing distance from the reference boundary [ $F(2,26) = 8.87, p = .001$ ]. In both one-way tasks, on the other hand, standard stimuli were classified as quickly as the other two stimuli in the category (both  $F$ s  $< 1$ ).

## Discussion

Experiment 4 demonstrates how classification of one HARD dimension is affected by dual context established by another HARD dimension. Task performance for pairs of HARD dimensions, here pitch and timbre, mimicked the pattern found when each of these dimensions was paired with loudness, a SOFT dimension. Values on HARD dimensions were classified in one-way tasks as fast as or faster than they were discriminated in standard-baseline tasks. Thus, as we found when pairing a HARD with a SOFT dimension, subjects can focus on values of pitch or timbre in the face of intraclass context when redundant context is present simultaneously, and, indeed, they may even show a gain relative to baseline performance. The pattern was the same throughout this experiment, regardless of whether the category contained two or three stimuli.

Classification of pitch and timbre in dual context is very different from classification of these same dimensions in Garner's single context. Melara and Marks (1989) showed that, relative to standard-baseline performance, intraclass context causes substantial Garner interference when classifying either pitch or timbre. In contrast, we show here that neither pitch classification nor timbre classification suffers when in dual context. Thus, Experiment 4 serves as a striking example of the different processes tapped by Garner's paradigm, on the one hand, and our new multiclass paradigm, on the other.

Pitch and timbre matched for baseline discriminability. Given this, it is interesting that gains from redundant context were stronger when subjects classified according to timbre than when they classified according to pitch. And, in a similar way, it is interesting that RTs for standard stimuli differed greatly depending on whether values of pitch or timbre were classified: For one-way timbre judgments, subjects were as fast at classifying standard stimuli as at discriminating them (in the correlated-baseline task), but, for one-way pitch judgments, standard stimuli were classified much slower than they were discriminated. These results indicate that, when embedded in classification tasks, values of timbre, at least of the sort used here, are somewhat more salient to subjects than values of pitch. In Experiment 5, we will examine how enhancing the baseline discriminability of pitch affects performance in the multiclass procedure.

As we have found in previous experiments, performance in two-way tasks fell intermediate between the slower one-way pitch performance and the faster one-way timbre performance. Perhaps classification in two-way tasks is achieved by pooling information from the second dimension. Alternatively, perhaps subjects focused on timbre (the more salient dimension) but suffered some loss in efficiency because values within each category varied. Unfortunately, the present results do not permit us to decide between these (or other) possibilities.

The most intriguing results of this experiment emerged when we compared speeds to classify individual stimuli within a category. In previous experiments, when a HARD dimension was paired with a SOFT one, we consistently found an effect of distance from the reference boundary: The closer a stimulus sat to that boundary, the longer subjects took to classify it. This experiment, however, in which two HARD dimensions were paired, revealed virtually no effect of boundary distance. By and large, standard stimuli were classified as quickly as other, more remote stimuli in the category. This null finding held in both one-way pitch and one-way timbre tasks for both two-stimulus and three-stimulus categories. Yet, interestingly, two-way tasks did show the familiar distance effects.

The absence of distance effects in one-way tasks perhaps reflects important differences between classifying attributes when both dimensions are HARD and classifying attributes when one dimension is HARD and the other SOFT. These differences may be based in optional strategies that subjects choose in the course of classifying, or they may have a more fundamental, mandatory origin in information processing. In either case, it seems clear that simple distance models of similarity are poorly equipped to handle such differences. On the one hand, the presence of such distance effects for HARD-SOFT combinations fits well with these models, since the physical (and presumably the psychological) dissimilarity of stimuli serves as a good predictor of RT performance. On the other hand, the absence of such distance effects for HARD-HARD combinations poses an obvious problem for these models, because

it is unclear why physical dissimilarity should not matter in this case. We address this issue further in the next experiment.

### EXPERIMENT 5

In Experiment 5, we had two goals. First, we sought to replicate the set of effects uncovered in Experiment 4. Specifically, we were interested in whether both pitch and timbre display resistance to intraclass context when redundant context is also present. Also, we wished to replicate, with a new group of subjects, the absence of distance effects in one-way classification tasks.

Our second goal was to examine how the mismatching of baseline discriminabilities affects performance. We used greater differences in frequency here but the same values of duty cycle. Our rationale was that by increasing the discriminability of the pitch dimension, we might lessen the interference from intraclass variation of the timbre dimension (see Felfoldy & Garner, 1971). Under such conditions, we may then ask whether pitch still yields noticeably smaller effects in the multiclass procedure.

#### Method

**Subjects.** Fourteen subjects (5 men and 9 women) were paid \$5 per hour to participate.

**Stimuli, Apparatus, and Procedure.** All aspects of Experiment 4 were duplicated, with the exception of the values of pitch used (see Figure 5). In Category A, stimuli were 860, 880, or 900 Hz. In Category B, stimuli were 940, 960, or 980 Hz.

#### Results

**Overall performance.** RTs and error rates appear in Table 9. The overall error rate was .05; errors correlated .77 with latency. As expected, subjects judged pitch at baseline (291 msec) significantly faster than they judged timbre at baseline [321 msec;  $F(1,13) = 12.9, p < .01$ ].

**Analyses involving tasks.** Speed to classify attributes in one-way pitch tasks (290 msec) was equal to pitch discrimination at baseline [291 msec;  $F(1,13) = .16, n.s.$ ]. This finding was consistent for two-stimulus categories and three-stimulus categories analyzed separately.

Speed to classify attributes in one-way timbre tasks (302 msec) was significantly faster than baseline [321 msec;  $F(1,13) = 5.77, p = .03$ ]. This difference was reliable for two-stimulus categories [297 msec;

$F(1,13) = 9.34, p < .01$ ], but not for three-stimulus categories [314 msec;  $F(1,13) = .41, n.s.$ ].

Performance in two-way tasks (305 msec) was no different from overall baseline performance [306 msec;  $F(1,13) = .01, n.s.$ ]. This finding was consistent for both two-stimulus and three-stimulus categories. ANOVA showed that one-way pitch judgments were made marginally faster than either one-way timbre judgments or two-way judgments [ $F(2,26) = 3.01, p = .07$ ]. There was no difference between overall performance in one-way tasks (296 msec) and two-way tasks [305 msec;  $F(1,13) = 2.33, n.s.$ ].

As before, a category-width effect obtained here, with performance in small-width tasks (289 msec) edging that in large-width tasks [301 msec;  $F(1,13) = 8.43, p = .01$ ].

**Analyses involving stimuli.** RTs to individual stimuli appear in Table 10. Speed to discriminate standard stimuli in the correlated-baseline task (270 msec) was compared with speeds to classify standard stimuli in one-way pitch tasks (281 msec), one-way timbre tasks (300 msec), or two-way tasks (310 msec). In each case, standard stimuli were discriminated significantly faster than they were classified [one-way pitch,  $F(1,13) = 8.67, p = .01$ ; one-way timbre,  $F(1,13) = 15.03, p < .01$ ; two-way,  $F(1,13) = 14.82, p < .01$ ]. Among these three types of classification tasks, speed to classify standard stimuli differed significantly [ $F(1,13) = 7.83, p < .01$ ], with that in one-way pitch tasks being faster than that in either of the other two tasks.

In two-stimulus categories, standard stimuli (293 msec) were classified 5 msec faster than distant stimuli [298 msec;  $F(1,13) = 4.49, p = .05$ ]. In three-stimulus categories, there was no difference overall among standard, adjacent, and distant stimuli [ $F(2,26) = .44, n.s.$ ]. However, RTs to these three stimuli interacted with the type of classification task [ $F(4,52) = 3.95, p < .01$ ]. In one-way pitch tasks, subjects responded to standard stimuli (281 msec) significantly faster than to either adjacent stimuli (298 msec) or distant stimuli [296 msec;  $F(2,26) = 7.88, p < .01$ ]. In two-way tasks, subjects responded to standard stimuli (320 msec) slower than to either adjacent stimuli (313 msec) or distant stimuli (312 msec), although these differences were not reliable [ $F(2,26) = 1.63, n.s.$ ]. Subjects responded to the three

Table 9  
Mean Reaction Times (in milliseconds) to Perform One-Way Pitch,  
One-Way Timbre, and Two-Way Tasks (Experiment 5)

Category (Task)	One-Way				Two-Way		Overall	
	Pitch		Timbre		Proportion		Proportion	
	<i>M</i>	Proportion of Error	<i>M</i>	Proportion of Error	<i>M</i>	Proportion of Error	<i>M</i>	Proportion of Error
Standard baseline	291	.05	321	.07			306	.06
Correlated baseline							270	.03
Two-stimulus (small width)	285	.03	293	.03	289	.04	289	.03
Two-stimulus (large width)	292	.03	300	.05	313	.06	301	.05
Three-stimulus	292	.05	314	.05	315	.06	307	.05

**Table 10**  
Mean Reaction Times (in milliseconds) to Stimuli in One-Way Pitch, One-Way Timbre, and Two-Way Tasks (Experiment 5)

Stimuli	One-Way		Two-Way	Overall
	Pitch	Timbre		
Two-Stimulus Category				
Standard	281	300	310	293
Other	297	307	290	298
Three-Stimulus Category				
Standard	281	315	320	306
Adjacent	298	313	313	308
Distant	296	313	312	308

types of stimuli equally fast in the one-way timbre tasks. These findings are presented graphically in Figure 6.

### Discussion

Experiment 5 replicated Experiment 4 in all important respects. Neither one-way pitch judgments nor one-way timbre judgments were performed slower than standard-baseline measures. Thus, both dimensions fit our notion of a HARD dimension. And, as in Experiment 4, timbre showed a gain in one-way classification (relative to baseline), whereas pitch did not. Thus, despite the fact that here pitch was much more discriminable than timbre, subjects were less able to take advantage of the correlated component when performing this type of classification. The fact that subjects still showed a 21-msec redundancy gain in the correlated-baseline relative to standard-baseline judgments of pitch [270 vs. 291 msec;  $F(1,13) = 22.75$ ,  $p < .001$ ] argues against the interpretation that pitch judgments have reached a performance ceiling.<sup>2</sup>

These findings suggest that timbre is intrinsically more salient under conditions of dual context. In a sense, then, timbre would appear to be a "HARDER" dimension than pitch. This interpretation is consistent with the gain we

found in one-way timbre judgments when the unattended dimension was loudness (Experiment 3). In each of the experiments involving classification of timbre, subjects not only avoided interference from intraclass context when classifying timbre, but were extremely successful at capitalizing on redundant context. It will be important in future study to discriminate what characteristics of spectral timbre permit subjects to focus so effectively on this dimension in dual-context classification.

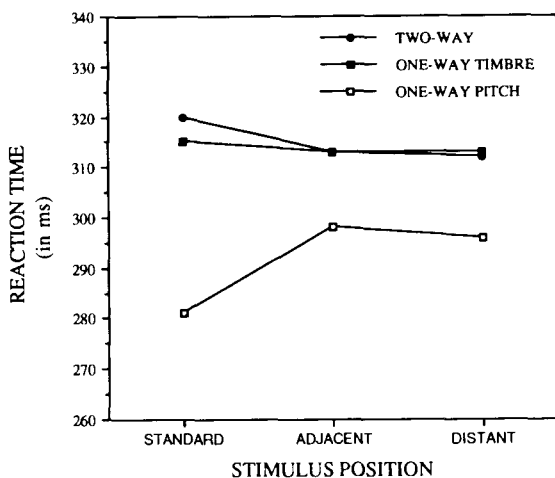
The most striking finding from Experiment 5 was the interaction between stimulus position (in three-stimulus categories) and the type of classification. In two-way tasks, standard stimuli were classified less well than adjacent stimuli or distant stimuli (see Figure 6), a finding obtained in Experiment 4 and throughout this study (e.g., see Figure 4). In one-way pitch tasks, however, standard stimuli were actually classified faster than either adjacent or distant stimuli, despite their proximity to the reference boundary (see Figure 6). And, as in Experiment 4, in one-way timbre classifications, stimuli in the standard position were classified as quickly as stimuli in either of the other two positions in the category (see Figure 6).

Clearly, then, a simple distance-to-boundary model is inadequate to account for speed to classify stimuli. Our findings indicate that in classification, two factors are especially relevant. First, speed depends on what classification rule is used. The fact that distance-to-boundary has characterized all two-way classifications in this study but only some of the one-way classifications argues that processing differs importantly between one-way and two-way classification rules. Second, speed depends on the HARDNESS of each member of the pair of interacting dimensions used to construct the stimulus set. Distance-to-boundary was useful when a HARD and a SOFT dimension were paired, but not when two HARD dimensions were paired.

### GENERAL DISCUSSION

The present experiments demonstrate convincingly that our approach of combining contexts reveals differences among interacting dimensions that were previously invisible in the single context of Garner's speeded classification procedure. As one example, the dimensions pitch and loudness display equivalent effects in Garner's procedure: Both dimensions show interference from intraclass context and gain from redundant context. In the multiclass paradigm, however, pitch and loudness behave differently: Pitch is HARD in the face of dual context, but loudness is SOFT. Thus the present study demonstrates the merit of using the original Garner paradigm in tandem with the new multiclass paradigm. The two paradigms provide different but compatible information regarding the processing of interacting dimensions.

For the dimensions that we have studied, it appears that characteristics of the attended dimension exert pre-eminence in processing over characteristics of the unattended dimension. Pitch was paired with loudness in Ex-



**Figure 6.** Mean reaction time (in milliseconds) to classify standard, adjacent, and distant stimuli in one-way pitch, one-way timbre, and two-way tasks (Experiment 5).

periments 1A and 2, and paired with timbre in Experiments 4 and 5. The pattern of RTs we found for one-way pitch classifications was comparable for both dimensions, indicating that dual context had similar effects on perceived pitch, whether the unattended dimension was *HARD* or *SOFT*. Similarly, attributes of timbre showed a gain in one-way classification, whether timbre was paired with loudness or with pitch. These results underscore a recurrent theme in this paper: Processing of the stimulus whole is not solely the effect of redundant or intraclass context on attribute perception, but involves an interaction between these contexts and intrinsic characteristics of stimulus dimensions. Context constrains perception, but only to a point. That is, the effects of dual context on the perception of attributes from an attended dimension are consistent, no matter what unattended dimension is used to create that context.

We note that before the foregoing conclusion can be made with certainty, a final pair of interacting dimensions, namely, the *SOFT-SOFT* pair, must be tested in the complete multiclass procedure. One possible *SOFT-SOFT* pair might be the color dimensions saturation and brightness. The question is this: Does the pattern we found for loudness still obtain when the unattended dimension is also *SOFT*? By testing color dimensions, we not only address this question, but we also evaluate whether the concept of *HARD-SOFT* is applicable to dimensions from other sensory modalities. Thus, we are currently testing multiclass performance of color dimensions.

### Characterizing Dimensional Interaction

What processes underlie interaction between a pair of dimensions? Melara and Marks (in press) recently proposed a model of dimensional interaction which holds that the perception of interacting dimensions involves processing two qualitatively different sources of information. *Attribute-level* information is the activation of representations corresponding to values on dimensions. Attribute-level information is extracted from continua, like pitch and loudness, that (as indicated by other criteria) are meaningful psychologically. Presumably, the perceiver can access attribute-level information whether two dimensions are interacting or separable. *Stimulus-level* information, on the other hand, is, in our view, limited to interacting dimensions. Stimulus-level processing joins attributes of pitch and loudness into a coherent whole (i.e., sound). Metaphorically, stimulus-level processing involves perception of the glue binding attributes together.

Our model contrasts with other, more traditional views of integrality, which assume that interaction entails, initially, holistic processing in the absence of dimensional information (e.g., Garner, 1974, 1981; Lockhead, 1972, 1979; Kemler, 1983a, 1983b; Smith & Kemler, 1977, 1978). According to these traditional accounts, what we call attribute-level information (e.g., specific pitches or loudnesses) is available only in a later, optional stage of processing. Melara and Marks (in press) have demonstrated, however, that subjects mandatorily process

attribute-level information from auditory stimuli according to specific, psychologically "real" orientations of stimulus axes—namely, pitch, loudness, and timbre (see also Grau & Kemler Nelson, 1988). Our data indicate, therefore, that attribute-level processing is as fundamental as the processing of information at the level of the stimulus whole.

### The Processing of Stimulus-Level Information

In our view, stimulus-level information is the contextual influence of one attribute on another. In other words, the metaphoric gluing of two attributes from a pair of interacting dimensions is, in reality, how context affects the perceptual meaning of those attributes; with pairs of separable dimensions, there is no glue, and hence no change in meaning.

The influence of context is evident in Garner's paradigm. Here stimulus-level processing may reveal itself in either of two ways: Garner interference from orthogonally varied dimensions, or redundancy gain from correlated dimensions. In our terminology, Garner interference arises when information at the level of the stimulus whole affects the processing of attributes. In filtering dimensions of sound, for example, subjects are presumably aware of (can attend separately to) the pitch and loudness attributes that constitute a given sound. But the perceptual experience of any one attribute depends on the other attribute to which it is attached. The attribute *high pitch* paired with the attribute *loud* differs phenomenally from *high pitch* paired with *soft*. Context established by the presence of loudness attributes acts to weight the perception of pitch attributes. In our view, then, filtering involves a classification of attributes, and not of stimuli, where the attributes in any one category vary in shade of meaning. Garner interference is due to the fact that there is greater within-category shifting (there is a many-to-one mapping of attributes on response keys) in the filtering task relative to the baseline task (in which there is a one-to-one mapping; see Garner, 1974; Lockhead, 1979; Lockhead, Gruenewald, & King, 1978).

Stimulus-level information enhances performance when dimensions are correlated. We claim that this effect is also due to the contextual changes surrounding attributes. Consider Figure 1. When the two pitches of Stimuli A and B are discriminated in a correlated-dimensions task, perceptual context varies with the attributes (i.e., low pitch co-occurs with soft; high pitch co-occurs with loud). In our view, by varying context, the pitches are actually more dissimilar perceptually (and, hence, better discriminated) when covaried with loudnesses than when loudness is held constant. Redundancy gain is thus attributable to the heightened discriminability afforded attributes in separate contexts.

Our emphasis on how context affects attributes again contrasts with traditional accounts of dimensional interaction. According to those views, subjects base identifications on the stimulus whole or "blob" (Lockhead, 1972, 1979). Thus attribute-level information need not be ac-



cessed (and is not available automatically) in order to perform the task. Alternatively, we claim that subjects are fully capable of basing their responses on attributes constituting stimuli—specifically, on values of the criterial dimension. In this sense, then, there is no blob. We argue instead that, with interacting pairs of dimensions, stimulus-level or contextual information is processed, and that it mandatorily affects (constrains) how those attributes are perceived.

### Characterizing Dimensional HARDNESS

The present study allows us to extend our view of dimensional interaction. Here we have explored how stimulus-level processing affects access to different types of attribute-level information. We delineated two kinds of contextual information that might affect the processing of attributes: redundant context, which leads to better performance due to greater perceptual discriminability, and intraclass context, which leads to poorer performance due to greater within-category shifting. As an approach to dissociating redundant from intraclass context, we suggested that different interacting dimensions might be processed differently were these two forms of context to be present simultaneously in a classification task. Such dual-context classifications (i.e., multiclass) allowed us to uncover previously unknown properties of interacting dimensions, namely, whether a given dimension is **HARD** or **SOFT**.

What does it mean for a dimension to be **HARD** or **SOFT**? Thus far, we have defined **HARD** and **SOFT** only operationally: **SOFT** dimensions show interference in one-way classification relative to baseline, whereas **HARD** dimensions show no interference. And, thus far, even these operational definitions have been realized only with simple auditory dimensions. The question remains, What is it about these dimensions that causes such differences in RT performance? At present we cannot answer this question. Nonetheless, while restricting our speculations to the auditory dimensions tested here, we now discuss how best to conceive of dimensional **HARDNESS**.

Consider the **HARD** dimension pitch. Attributes along this dimension differ in continuous steps that, psychophysically, follow closely the physical dimension frequency. Thus, in terms of an underlying representational code, perceivers can represent differences in pitch as differences in degree along some low-high continuum. Call this coding scheme a *continuum code*. When pitch is filtered in Garner's task, only intraclass context is present, meaning that distance between categories is unidimensional—that is, provided only by differences along the criterial dimension (see Figure 2, panel A). In such cases, we suggest that subjects may base classifications solely on values of the continuum code. That is, subjects may make greater-than/lesser-than judgments of points along a continuous dimension. And, in line with our model of stimulus-level processing, intraclass varia-

bility causes shifts along this continuum that lead, in turn, to decision interference.

When pitches are classified in the multiclass procedure, redundant context is added to the intraclass context of the filtering task. Adding redundant context establishes a between-category distance that is two-dimensional: Categories are separated along both the horizontal and the vertical axes (see Figure 2, panel B). We suggest that for pitch (or timbre), this two-dimensional separation may permit processing of attributes in terms of kind rather than merely in terms of degree. In other words, dual context may permit access to a different scheme of coding pitches or timbres, what we will call a *kind code*, which represents attributes discretely rather than continuously. For example, in the simplest case, a kind code might be the arithmetic average of a set of continuous values. Such coding is perhaps linguistically based, but, in any case, it is certainly derived later than continuum coding.

Why does a two-dimensional (dual context) separation entail access to kind coding of pitch or timbre, whereas a one-dimensional (intraclass context) separation does not? One possibility is that dual context permits criterial attributes to be "contained" apart. The second dimension, through its correlation with the criterial dimension, encourages erection of a perceptual boundary that segregates attributes in Category A from attributes in Category B. This boundary may or may not be oriented like (or have the form of) our reference boundary (see Figure 3). Nonetheless, like the reference boundary, such a perceptual boundary would provide attributes in one category with an identity qualitatively independent of attributes in the other category. Kind coding, therefore, is a more absolute coding scheme, and entails a severing of the metaphorical links that connect attributes in the relativistic continuum scheme. Dual context, by holding apart attributes on **HARD** dimensions, permits access to the absolute code.

Are **SOFT** dimensions coded by kind? It is conceivable that they are, implying that **HARD** and **SOFT** dimensions differ regarding the precision with which continuum codes map onto kind codes. With **HARD** dimensions, the mapping might have a high resolution, perhaps corresponding closely to just-noticeable differences (JNDs). With **SOFT** dimensions, on the other hand, the resolution might be comparatively low, so that attributes on these dimensions must be highly discriminable from each other before classification can benefit from the kind code. In the present study, we used perceptually small differences between values on dimensions. To test this hypothesis, therefore, it will be necessary in future research to compare multiclass classification of **SOFT** attributes at several levels of baseline discriminability.

Why do **HARD** and **SOFT** dimensions differ from each other regarding either the existence of, or speed of access to, kind codes? This is theoretically perhaps the most fundamental question, but one for which, at present, we

can offer no satisfactory answer. One possibility is that certain differences exist at the neural level, perhaps akin to those suggested by Stevens (1957) in distinguishing metathetic from prosthetic continua. An alternative possibility is that experientially based priorities are assigned by the perceptual system for processing information from different dimensions. The auditory dimensions that we call HARD might be those that, over the course of development, were most useful in recognizing a sound or a group of sounds. Because of their long-term usefulness, such dimensions may be accorded privileged access to discrete codes. This hypothesis is attractive, because it suggests an empirically testable developmental course for the emergence of dimensional HARDNESS. The hypothesis is reminiscent of that offered by Kemler Nelson and her colleagues (Kemler, 1983a, 1983b; Smith & Kemler, 1977, 1978) to explain a perceptual shift found with certain pairs of dimensions (e.g., size and brightness) that behave like integral dimensions in children, but like separable dimensions in adults. Given Kemler Nelson's model, it would be interesting to know whether such pairs of dimensions are also perceived as HARD by children. Similarly, it would be important to examine whether dimensions that are HARD for adults are as HARD for children. A developmental hypothesis suggests that they would not be: HARD dimensions may be conferred their privileged access only after years of faithful service.

### Dimensional HARDNESS Versus Dimensional Separability

Although we have restricted reference of HARD and SOFT to interacting dimensions, it is tempting, especially given a developmental hypothesis, to connect dimensional HARDNESS with dimensional separability. For both HARD dimensions and separable dimensions, subjects are able, relatively speaking, to focus attention on one of two attributes in the face of trial-to-trial changes in the stimuli presented. However, we believe that the analogy of HARD and separable may be a misleading one. The ability to focus on attributes from a separable dimension results from the failure of the context established by the other dimension to affect perceptual processing; with separable dimensions, the absence of any stimulus-level information is the critical factor. On the other hand, the ability to focus on attributes from a HARD dimension presumably arises through the perceptual context created by the other stimulus dimension; with HARD dimensions, it is the presence of certain kinds of stimulus-level information—namely, redundancy information in a classification task supplied by correlating interacting dimensions—that appears critical in permitting subjects to focus on attributes. Thus we believe that the similarities between HARD and separable dimensions may be superficial and overshadowed by substantive differences. Nevertheless, how HARD dimensions differ from separable dimensions is an empirical question, one that can be studied profitably in future research

that tests how separable dimensions themselves behave in the dual context of multiclass.

### REFERENCES

- ASHBY, F. G., & GOTT, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 33-53.
- ASHBY, F. G., & PERRIN, N. A. (1988). Toward a unified theory of similarity and recognition. *Psychological Review*, **95**, 124-150.
- FELFOLDY, G. L., & GARNER, W. R. (1971). The effects on speeded classification of implicit and explicit instructions regarding redundant dimensions. *Perception & Psychophysics*, **9**, 289-292.
- FLETCHER, H., & MUNSON, W. A. (1933). Loudness, its definition, measurement and calculation. *Journal of the Acoustical Society of America*, **5**, 82-108.
- GARNER, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.
- GARNER, W. R. (1976). Interaction of stimulus dimensions in concept and choice processes. *Cognitive Psychology*, **8**, 98-123.
- GARNER, W. R. (1981). The analysis of unanalyzed perceptions. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 119-139). Hillsdale, NJ: Erlbaum.
- GARNER, W. R., & FELFOLDY, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, **1**, 225-241.
- GARNER, W. R., HAKE, H. W., & ERIKSEN, C. W. (1956). Operationism and the concept of perception. *Psychological Review*, **63**, 49-159.
- GRAU, J. W., & KEMLER NELSON, D. G. (1988). The distinction between integral and separable dimensions: Evidence for the integrality of pitch and loudness. *Journal of Experimental Psychology: General*, **117**, 347-370.
- IMAI, S., & GARNER, W. R. (1965). Discriminability and preference for attributes in free and constrained classification. *Journal of Experimental Psychology*, **69**, 596-608.
- KEMLER, D. G. (1983a). Exploring and reexploring issues of integrality, perceptual sensitivity, and dimensional salience. *Journal of Experimental Child Psychology*, **36**, 365-379.
- KEMLER, D. G. (1983b). Holistic and analytic modes in perceptual and cognitive development. In T. Tighe & B. E. Shepp (Eds.), *Perception, cognition, and development: Interactional analyses* (pp. 77-102). Hillsdale, NJ: Erlbaum.
- KRUMHANS, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review*, **85**, 445-463.
- KRUSKAL, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, **29**, 1-27.
- LOCKHEAD, G. R. (1972). Processing dimensional stimuli: A note. *Psychological Review*, **79**, 410-419.
- LOCKHEAD, G. R. (1979). Holistic versus analytic process models: A reply. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 746-755.
- LOCKHEAD, G. R., GRUENEWALD, P., & KING, M. (1978). Holistic vs. attribute repetition effects in classifying stimuli. *Memory & Cognition*, **6**, 438-445.
- MARTIN, R. C., & CARAMAZZA, A. (1980). Classification in well-defined and ill-defined categories: Evidence for common processing strategies. *Journal of Experimental Psychology: General*, **109**, 320-353.
- MEDIN, D. L., WATTENMAKER, W. D., & HAMPSON, S. E. (1987). Family resemblance, concept cohesiveness, and category construction. *Cognitive Psychology*, **19**, 242-279.
- MELARA, R. D. (1989a). Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 69-79.
- MELARA, R. D. (1989b). Similarity relations among synesthetic stimuli and their attributes. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 212-231.

- MELARA, R. D., & MARKS, L. E. (1989). *Interaction between auditory dimensions: Timbre, pitch, and loudness*. Manuscript submitted for publication.
- MELARA, R. D., & MARKS, L. E. (in press). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception & Performance*.
- MELARA, R. D., & O'BRIEN, T. P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, **116**, 323-336.
- PODGORNY, P., & GARNER, W. R. (1979). Reaction time as a measure of inter- and intraobject visual similarity: Letters of the alphabet. *Perception & Psychophysics*, **26**, 37-52.
- POMERANTZ, J. R. (1983). Global and local precedence: Selective attention in form and motion perception. *Journal of Experimental Psychology: General*, **112**, 515-540.
- POMERANTZ, J. R. (1986). Visual form perception: An overview. In E. C. Schwab & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Visual perception* (pp. 1-30). New York: Academic Press.
- POMERANTZ, J. R., & GARNER, W. R. (1973). Stimulus configuration in selective attention tasks. *Perception & Psychophysics*, **14**, 565-569.
- PRICE, J. M. (1979). Software timing for 6500 series microcomputers. *Behavior Research Methods & Instrumentation*, **11**, 568-571.
- SHEPARD, R. N. (1962a). The analysis of proximities: Multidimensional scaling with an unknown distance function I. *Psychometrika*, **27**, 125-140.
- SHEPARD, R. N. (1962b). The analysis of proximities: Multidimensional scaling with an unknown distance function II. *Psychometrika*, **27**, 219-246.
- SMITH, L. B., & KEMLER, D. G. (1977). Developmental trends in free classification: Evidence for a new conceptualization of perceptual development. *Journal of Experimental Child Psychology*, **24**, 279-298.
- SMITH, L. B., & KEMLER, D. G. (1978). Levels of experienced dimensionality in children and adults. *Cognitive Psychology*, **10**, 502-532.
- STEVENS, S. S. (1957). On the psychophysical law. *Psychological Review*, **64**, 153-181.
- TORGERSON, W. S. (1958). *Theory and method of scaling*. New York: Wiley.
- WOOD, C. C. (1975). Auditory and phonetic levels of processing in speech perception: Neurophysiological and information-processing analyses. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 3-20.

## NOTES

1. Our reference boundary represents only one of many possible decision rules that subjects might invoke in a classification task (see, e.g., Ashby & Gott, 1988; Ashby & Perrin, 1988). We do not assume that subjects necessarily use such a boundary to make classification decisions. Rather, the boundary will serve as a convenient standard for us when making comparisons among RTs to various stimuli.

2. We thank an anonymous reviewer for suggesting this interpretation.

(Manuscript received June 26, 1989;  
revision accepted for publication October 11, 1989.)