Journal of Experimental Psychology: Applied
2000, Vol. 6, No. 1, 15–30

Copyright 2000 by the American Psychological Association, Inc.
0278-7393/00/$5.00 DOI: 10.1037//0278-7393.6.1.15

Pitch and Pitch Change Interact in Auditory Displays

Bruce N. Walker
Rice University

Addie Ehrenstein
Utrecht University

Designing auditory displays requires understanding how different attributes of sound are processed. Operators must often listen to a particular stimulus dimension and make control actions contingent on the auditory information. Three experiments used a selective-listening paradigm to examine interactions between auditory dimensions. Participants were instructed to attend to either relative pitch or direction of pitch change of dynamic stimuli. With vertically arranged keypress responses, reactions to both dimensions showed stimulus–response compatibility effects, indicating that pitch is treated spatially. Direction of pitch change affected responses to pitch; level of pitch more strongly affected responses to pitch change. To reduce deleterious effects of irrelevant pitch information, auditory display designers can restrict the pitch range used to display dynamic data.

The performance of many types of tasks in a variety of real-world settings requires selective attention to just one dimension of a stimulus. For example, to avoid a collision, a pilot might need to make a speeded response on the basis of the proximity of an oncoming plane, ignoring, for the moment at least, the size or color of it. Often, the displayed information is auditory rather than visual, as it is for the Geiger counter operator who may need to listen selectively to the temporal pattern of a sound, which indicates the prevalence of radioactive particles, and ignore changes in the pitch of the sound, which may indicate the type of particles that are present. The increasing use of auditory displays means that a growing number of professionals rely on sounds emitted in their environments, by their tools, or from communications devices to guide their actions. To design auditory interfaces that afford better comprehension and elicit faster and more accurate reactions, it is necessary to understand how different attributes of auditory stimuli interact to influence perception and responding.

When the correctness of a response depends on selectively attending to one dimension of a sound—and ignoring other stimulus dimensions—performance depends on how the individual dimensions of the sound are perceived and how different dimensions combine to influence performance. It is well known that the perception of one aspect of a sound (e.g., loudness) depends not only on the physical characteristic primarily associated with that attribute (amplitude, in the case of loudness) but also on other characteristics not generally associated with it (e.g., frequency). For example, Robinson and Dadson (1957) documented the equal-loudness contours that reflect how the perceived loudness of two tones of equal amplitude can differ, depending on the frequencies of the sounds. Moreover, using above-threshold sounds, Neuhoff and McBeath (1996) found that judgments of loudness are affected by changes in frequency. Sometimes, different dimensions of auditory stimuli interact in nonobvious ways. For example, Melara and Marks (1990b) showed that listeners attending to the timbre of sounds responded faster to a hollow sound if the sound was also high in pitch, rather than low, and responded faster to twangy sounds when the pitch was low, rather than high—despite instructions to ignore pitch. Pitch judgments are affected by variations in timbre in a similar manner (Melara & Marks, 1990b).

Of the various attributes of sound that can be used to display information, it can be argued that pitch is of primary practical interest because currently, it is the dimension most commonly used to represent data in auditory displays, both for the relative ease with which pitch can be controlled by current display hardware (Kramer, 1994) and because the basics of pitch perception are well documented (e.g., Deutsch, 1982; Moore, 1989; Stevens, 1957). Other benefits of using pitch as a display dimension are that most listeners are familiar with the concept of pitch and can detect fairly small pitch changes with little training (e.g., a less than 5-Hz change in a pure tone of 1000 Hz; Wier, Jesteadt, & Green, 1977) and that pitch more evenly represents a wider range of values than other perceived dimensions of sound, such as loudness, which at the extremes, does not provide an effective display dimension (soft sounds are masked by ambient noise, and loud sounds are potentially disturbing or even damaging; e.g., Patterson, 1982).

Because data to be represented in an auditory display rarely consist of single, static values (which would result in the presentation of unchanging pitches), it is important to know how the classification of pitch is affected by dynamic changes in the stimulus, such as changes in pitch itself. Everyday applications that rely on the interpretation of dynamic data range from storm tracking to fetal heart monitoring (see Neuhoff & McBeath, 1996, for more

---

Bruce N. Walker, Department of Psychology, Rice University; Addie Ehrenstein, Department of Psychonomics, Utrecht University, Utrecht, The Netherlands.

Correspondence concerning this article should be addressed to Bruce N. Walker, Department of Psychology, Rice University, 6100 Main Street, Houston, Texas 77005-1892, or to Addie Ehrenstein, Department of Psychonomics, Heidelberglaan 2-17, 3584 CS Utrecht, The Netherlands. Electronic mail may be sent to walkerb@rice.edu.

examples). As mentioned earlier, Neuhoff and McBeath (1996) have investigated the relation between frequency and amplitude in pitch and loudness perception using dynamic auditory stimuli. They found that the subjective experience of specific dimensions of an auditory stimulus depends on whether the stimulus is static or dynamic. In particular, Neuhoff and McBeath studied what they call the "Doppler illusion," wherein perception of pitch is affected by dynamic changes in the loudness of a sound over the course of a trial. For example, the measured frequency of the sound of an oncoming train actually drops as the train approaches, yet listeners often report an increase in perceived pitch. When the loudness is held static, or even increased in steps, the illusion is no longer evident. Perception of loudness is similarly influenced by dynamic changes in pitch. The authors concluded that "dynamic intensity change can influence perceived pitch in a manner *opposite* [italics added] that specified by static intensity change" (Neuhoff & McBeath, 1996, p. 979). Few, if any, other studies have explored dimensional interaction with dynamic auditory stimuli.

The present research continues the investigation of stimulus interactions in dynamic auditory stimuli by exploring the effects of pitch change itself on the classification of pitch, as well as the effects of relative pitch on the classification of pitch change. In addition, by examining how the nature of the mapping of auditory stimuli to their assigned responses affects performance, we also explore the manner in which relative pitch is coded by the listener. In particular, we test whether pitch is coded as a spatial dimension (with higher pitches corresponding to higher positions) by looking for changes in performance resulting from the nature of the assignment of the stimuli to a vertically oriented response set. These questions about pitch are specific examples of two more general issues: the extent to which different dimensions of the stimulus interact (stimulus–stimulus interactions) and the extent to which characteristics of the stimuli interact with characteristics of the responses (stimulus–response effects). We address each of these issues in turn.

Interactions of Pitch With Other Stimulus Dimensions

Many tasks use auditory displays to aid the interpretation of raw data. Such "sonification" tasks (see Kramer, 1994) rely largely on perceptual and cognitive processing. Typically, the task is to use the information provided (e.g., an auditory representation of the data) to generate new information (e.g., an interpretation or analysis of the data). These tasks, such as using auditory displays for the analysis of turbulence in fluids (Blattner, Greenberg, & Kamegai, 1990), interpretation of topological structures (Axen & Choi, 1995), or identification of internal flaws in concrete bridges (Valenzuela, Sansalone, Krumhansl, & Streett, 1997), require assimilation of complex auditory information. Flowers and his colleagues (e.g., Flowers, Buhman, & Turnage, 1997; Flowers & Hauer, 1992, 1993, 1995) have studied how some aspects of auditorily presented graphs affect statistical inferences about data, and B. N. Walker and Kramer (1996) made initial investigations into data-to-sound mapping choices for a quality-assurance task in a simulated factory. However, few studies have examined how listeners are able to attend to just one dimension of an auditory display (representing one dimension of the data) and how the various sound attributes interact in these tasks.

There has been considerable research into the more basic perceptual aspects of how auditory dimensions interact. Although no study that we know of has addressed interactions of pitch and pitch change, several researchers have found interactions of pitch with other stimulus dimensions, some of which we have already mentioned. Most of the studies we review used the speeded-classification paradigm (Garner & Felfoldy, 1970) to determine whether two different dimensions of a stimulus can be processed independently of each other or whether the dimensions interact such that judgments on one dimension are affected by varying the value of the other. To test for interactions between stimulus dimensions, response times (RTs) and accuracy in conditions in which one dimension is held constant (i.e., baseline conditions) are compared with other conditions in which *both dimensions are varied. The increase in RT from the* baseline condition to the so-called filtering condition, in which the two stimulus dimensions are varied orthogonally, is termed *Garner interference* (Pomerantz, 1983, 1986) and indicates that processing of the two dimensions is not independent.

For our purposes, the measure of greatest interest is whether congruity effects, in which some pairings of stimulus values lead to faster performance than others, are obtained. For congruity effects to occur, there must be some meaningful or stereotypical relation between the values of the two stimulus dimensions (Clark & Brownell, 1976; Melara & O'Brien, 1987). Congruity effects are measured by the difference in RT between congruent trials, on which stimulus dimensions are paired such that the values of the dimensions correspond, and incongruent trials, on which they do not. Such effects have been found for a surprising variety of pairings of stimulus dimensions. One example, mentioned above, is that timbre interacts with pitch such that hollow sounds go with high pitch, whereas twangy sounds are classified more quickly when accompanied by low pitch (Melara & Marks, 1990b).

Melara and Marks (1990c) also found that varying loudness affects pitch judgments under conditions in which selective attention to pitch is encouraged. In this case, congruent trials are ones on which loud sounds are high in pitch and soft sounds are low in pitch. Cross-modal studies have demonstrated that pitch interacts with some nonauditory stimulus dimensions as well. Marks (1987; see also Schiller, 1935) found that judgments of pitch were affected by whether a simultaneously presented visual stimulus was bright or dim (where high pitch corresponds to bright), as well as by whether the accompanying visual stimulus was black or white (high pitch corresponds to white). Marks (1987) also found that judgments of sharp-edged, jagged shapes and of smooth, rounded shapes were affected by the pitch of an accompanying sound, such that sharp shapes corresponded with high pitch. Interactions between pitch and spatial attributes of an accompanying stimulus have been found in several contexts. One example is that the spatial location of a visual stimulus that is presented concurrently with a high- or low-pitched sound influences classification of that sound (e.g., responses to high pitches are faster when the location of the visual stimulus is relatively high; Melara & O'Brien, 1987). Even when spatial information is conveyed indirectly, such as through the presentation of a word, pitch judgments are affected. Melara and Marks (1990a) found that pitch classification depended on the identity of a concurrently presented word, whether displayed visually (*HI* or *LO*) or spoken (*high* or *low*). As might be expected on the basis of the findings of Melara and O'Brien, high pitch showed evidence of being congruent with the word *high* and low pitch with the word *low*. In all of these studies, not only were interactions of pitch with other stimulus dimensions found, but also particular pairings of the stimuli led to faster classification.

The findings that pitch interacts with the spatial position of visual stimuli and the content of spatial words, in particular, seem to have important implications for the display of auditory information. If pitch is represented as a spatial dimension, it is likely to interact with other spatial representations, such as representations of other data in the task environment or with the representations of the responses.

Interactions of Stimulus and Response Dimensions

Some types of tasks involving auditory displays not only require comprehension of the sonified information but also lead to or influence a related control action. Examples include auditory traffic collision avoidance systems in airplanes (e.g., Begault, Wenzel, Shrum, & Miller, 1996), auditory cues used to help computer users browse the World Wide Web (e.g., Albers, 1996), and various types of reactions made to auditory warnings (e.g., Hellier, Edworthy, & Dennis, 1993; Leung, Smith, Parker, & Martin, 1997; Patterson, 1982). An early, but certainly very germane, application used sound to display the critical information in a cockpit that is used to fly an airplane. After only 1 hr of practice, pilots were prepared to fly by auditory instruments alone (Forbes, 1946).

It is well established that both the choice of response set and the way in which stimuli are assigned to responses can significantly affect performance (e.g., Kornblum, Hasbroucq, & Osman, 1990; Wang & Proctor, 1996). For example, in the case where right- or left-hand responses are made to sounds presented to the right or left ear, stimulus–response (S-R) compatibility effects are found such that responses are faster if the hand used to respond corresponds to the ear of presentation (e.g., a left-hand response to a sound presented to the left ear) than if hand and ear do not correspond. It is important to note that these findings are for responses with the hands in the "normal" position—that is, with the left hand responding to the left of the right hand. When the hands are crossed, the ear–hand relation usually reverses (Roswarski & Proctor, in preparation). Such effects of spatial position are found even when the relevant response dimension (e.g., the pitch of the stimulus or the gender of the speaker's voice) is not related to the ear of presentation (see Simon, 1990, for a review).

Among the most widely studied S-R compatibility effects are those that result from the correspondence of the spatial locations of stimuli and the spatial locations of responses (Proctor & Reeve, 1990). As mentioned above, spatial position can exert strong influences on performance even when the position of the stimulus is nominally irrelevant to the classification of the stimulus. The effects of spatial position on the classification of nonspatial stimuli is called the *Simon effect*. The hallmark of this effect is faster responses to nonspatial, relevant stimulus dimensions (such as color or shape) when the position of the stimulus (e.g., right side of the display) corresponds to the location of the response (e.g., right-side response) than when it does not (see Lu & Proctor, 1995, for a review). When the stimulus conveys spatial information, such as when the task is to respond with a left key when the spoken word *left* is heard, while ignoring the ear in which the word is presented, the effects of the to-be-ignored spatial location of the stimulus (e.g., in which ear the stimulus is presented) are often called spatial Stroop effects (e.g., Dyer, 1972; White, 1969). Although much theoretical work has been conducted on the nature of S-R compatibility, Simon, and spatial Stroop effects, for our practical purposes, it suffices to note the circumstances under which such effects might be expected to appear.

S-R compatibility effects are often attributed to population stereotypes (Fitts & Deininger, 1954) to classify and respond to stimuli in a certain way (e.g., a natural tendency to respond to a word by reading it). The correspondence between a stimulus and response set is sometimes called *dimensional overlap* (Kornblum et al., 1990). Dimensional overlap, or the sharing of categorical attributes, is considered to be a prerequisite for S-R compatibility effects to occur. Findings of interactions of pitch with the position of a visual stimulus and with the presentation of the words *high* and *low* in the speeded-classification paradigm suggest that pitch is related semantically, or categorically, to spatial position. Other evidence that pitch is perceived as a vertical spatial dimension, comes from studies by Mudd (1963) and R. Walker (1987). Mudd studied the mental representation of pitch by having participants listen to pairs of sounds and then move a peg on a pegboard from the center position (which represented the first sound) to any of the other holes in a 19 × 19 matrix to indicate the relationship of the second sound to the first. Listeners tended to place the peg corresponding to higher pitched sounds higher on the pegboard, consistent with the hypothesis that listeners treat pitch as being correlated with vertical spatial position, with higher pitch corresponding to higher positions. It should be noted that there was also a correlation (albeit a weak one) between higher pitch and locations further to the right on the pegboard. R. Walker looked for correspondences between dimensions of sound and visual stimuli in a study in which participants from several different sociocultural groups matched visual symbols to sounds. The sounds varied along four auditory dimensions, and the visual stimuli varied along four visual dimensions. R. Walker found that participants "displayed a proclivity for choosing systematically a visual metaphor for each of the four acoustic parameters in the following manner: frequency was matched with vertical placement, amplitude with size, waveform with pattern, and duration with horizontal length" (p. 496). R. Walker also noted that musical training, and to a lesser extent culture and environment, affected the consistency of visual–auditory pairings.

Simon, Mewaldt, Acosta, and Hu (1976) more directly examined the pitch–space stereotype that higher pitches correspond to higher spatial positions and lower pitches to lower positions in a task in which up and down toggle-switch movements were assigned to stimuli that were high or low in pitch and emanated from an upper or lower speaker. The relevant stimulus dimension was pitch and was uncorrelated with the irrelevant dimension, speaker position. The assignment of pitch to responses was either compatible with the pitch–space stereotype (i.e., high pitch assigned to an upward toggle-switch movement and low pitch assigned to a downward movement) or incompatible (pitches assigned in the reverse order). Simon and his colleagues found effects of the S-R assignment, such that responses were faster with the compatible mapping. Moreover, when the assignment was compatible, effects of the irrelevant dimension were found such that when the position of the sound corresponded to the response (e.g., the sound was presented from the upper speaker when the correct response was an upward movement), responses were significantly faster. In other words, when the irrelevant spatial dimension of the stimulus (i.e., upper vs. lower speaker) was spatially compatible with the correct response, responses were faster than when the spatial location of the sound and the response did not correspond.

No effects of correspondence of speaker position were found when the assignment of pitch to responses was incompatible. This latter finding can perhaps be explained in terms of the operations performed on the pitch stimulus to select the response. If it is assumed that participants in the incompatible assignment condition suppressed the spatially corresponding response to the pitch of the stimuli, this may have also suppressed effects of the correspondence to speaker position. However, the lack of effects of correspondence in the incompatible condition might also be explained in terms of the longer RTs in the incompatible assignment conditions: Effects of the irrelevant sound position may have dissipated in the time taken to select the noncorresponding response (Hommel, 1994).

Another line of research relevant to our current purposes has examined the classification of dynamic visual stimuli. It has been found that judgments of the physical position of a visual target are influenced by the direction of motion of the target, even when the task is to respond to, for example, onset position while ignoring the direction of motion (e.g., Ehrenstein, 1994; Michaels, 1988, 1993; Proctor, Van Zandt, Lu, & Weeks, 1993). When the task is to attend selectively to the onset position of a visual target (i.e., whether a moving dot or a square appears on the left or the right side of a display), responses are typically faster if the onset position (e.g., left) is congruent with the direction of motion of the same visual target (e.g., leftward) than if the position and direction of movement are incongruent (e.g., the square appears on the left, but moves toward the right side of the display; Ehrenstein, 1994; Proctor et al., 1993). In most cases, the responses to be made in such tasks are left or right keypress responses or deflections of a joystick to the left or right. In such a task, the effects of the to-be-ignored stimulus dimension are usually explained in terms of spatial coding (e.g., Heister, Schroeder-Heister, & Ehrenstein, 1990; Lu & Proctor, 1995). In the example above, both onset position and direction of movement are assumed to be coded as left or right, as are the possible responses. The spatial coding theory is supported by the finding that essentially the same results are obtained if, instead of motion, a stationary arrow is used as the direction cue (Proctor et al., 1993). When both onset position and the direction of the arrow match the assigned compatible response, performance is better than when either the designated relevant or irrelevant dimension is incongruent with the response. It seems reasonable to expect that pitch and pitch change might produce similar congruity effects, such that if a sound is high in pitch, responding is faster if the sound becomes higher in pitch than if it becomes lower in pitch, and vice versa.

Experiment 1

If it is possible to attend selectively to the relative pitch of a sound in an auditory display while ignoring the direction of pitch change, then the time it takes to respond to whether the sound is high or low in pitch, and the accuracy of this response, should not depend on pitch change. That is, whether a high-frequency tone becomes lower or higher in frequency and a low-frequency tone becomes lower or higher in frequency should not affect performance. However, if the pitch-change information intrudes on the pitch decision, the speed and accuracy of responses to high pitches that become higher and to low pitches that become lower (i.e., congruent stimuli) should be better than for responses to high pitches that become lower and low pitches that become higher (i.e., incongruent stimuli). The same arguments can be made regarding the influence of relative pitch on pitch-change judgments.

To investigate interactions between the dimensions of onset pitch and direction of pitch change, we presented sounds that started at a given pitch and then became higher (i.e., changed to a higher frequency) or lower (i.e., changed to a lower frequency) in pitch. We instructed participants to listen to the sounds, attending selectively either to the onset pitch or to the direction of pitch change, and then to make a speeded classification according to whether the pitch was high or low or became higher or lower, respectively.

We used a set of 12 stimuli (see Figure 1 and Table 1). Through pilot testing, we found that using such a relatively large set of stimuli seems to prevent participants from learning to associate a particular response with a given stimulus and to encourage participants to listen analytically to either the onset pitch or direction of pitch change. This type of analytical listening is also more representative of the sorts of tasks involving auditory presentation of dynamic data. It is important to note that for stimuli classified as high in terms of onset pitch, both the onset and final pitches were

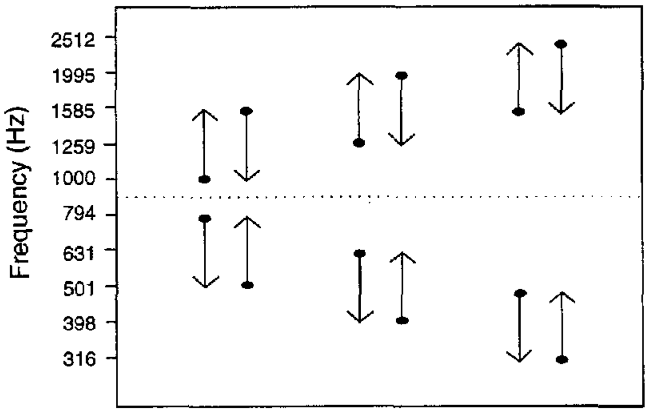

Figure 1. Schematic representation of the 12 auditory stimuli.

higher than any of the onset or final pitches of tones classified as low. That is, the high pitches did not overlap with the low pitches.

A second question of interest is how listeners code pitch. As discussed, previous evidence suggests that high pitches are coded as being high in a vertical sense—that is, as if they correspond to higher spatial locations than lower pitches. If this vertical coding of pitch is robust, then S-R compatibility effects would be expected to result from the assignment of pitch stimuli to a vertical response set, such that responding should be faster when high pitches (or upward movements of pitch) are assigned to an upper response key and low pitches (or downward movements of pitch) to a lower key than vice versa. To test this hypothesis, we also manipulated S-R assignment in this experiment.

Method

Participants. Twenty-eight Rice University undergraduates participated in the experiment for partial credit in a psychology course. All participants reported normal hearing, and none had participated in the pilot studies. Each participant completed two sessions, one in which onset pitch (i.e., the pitch at onset) was the relevant dimension and one in which pitch change (i.e., the direction of pitch change) was the relevant dimension. Presentation of the sessions was counterbalanced across participants. We assigned participants randomly to either a compatible or incompatible response assignment, with the constraint that there were equal numbers of participants in each assignment with each order of presentation of sessions.

Table 1
Initial and Final Frequencies of the Stimuli

	Onset		Offset	
Stimulus	Exponent	Frequency (Hz)	Exponent	Frequency (Hz)
1	3.2	1585	3.4	2512
2	3.1	1259	3.3	1995
3	3.0	1000	3.2	1585
4	3.4	2512	3.2	1585
5	3.3	1995	3.1	1259
6	3.2	1585	3.0	1000
7	2.7	501	2.9	794
8	2.6	398	2.8	631
9	2.5	316	2.7	501
10	2.9	794	2.7	501
11	2.8	631	2.6	398
12	2.7	501	2.5	316

Note. Stimuli were 150-ms pitch glides composed of 10 intermediate steps. The changes in the stimuli were equal in terms of log-frequency. The frequencies were calculated as $f = 10^N$, where f is the frequency of the sound and N is the exponent listed in the table. For example, Stimulus 3 has the onset frequency of $f = 10^{3.0} = 1000$ Hz.

Stimuli and apparatus. The stimuli were brief pitch glides composed of a series of short pitches (see Figure 1 and Table 1 for the initial and final frequencies of the stimuli). The total duration of each sound was 150 ms, with 10 intermediate steps creating an apparently continuous change in pitch. The changes in the stimulus pitch were made equal in terms of log frequency to equate the change in perceived pitch for all stimuli.

Stimuli for which the sound started high in pitch and became higher or started low and became lower in pitch were considered congruent because their onset pitch and direction of pitch change corresponded (the left stimulus in each group of two stimuli in Figure 1). The stimuli whose relative starting pitch and direction of pitch change were opposite were called incongruent (the right stimulus of each pair in Figure 1).

The stimuli can also be described in terms of their difference from the average pitch of the entire stimulus set, with each stimulus considered to have a small, medium, or large separation. The four sounds that were closest to the middle of the set were considered to have small separation (the four left-most stimuli in Figure 1). The four sounds that were highest and lowest relative to the middle of the pitch range had large separation, and the remaining four had medium separation. It is important to note that separation was defined not in terms of onset pitch but rather in terms of the average pitch for each stimulus.

We tested participants individually in a sound-attenuated testing room. An IBM-compatible 486-DX 33-MHz computer with a 14-in. (36-cm) color VGA monitor and standard 101-key keyboard running a program written in Micro Experimental Laboratory Professional 2.0 (MEL; Schneider, 1988, 1995) was used to present instructions and to control stimulus presentation and data collection. Stimuli were presented through the computer's internal speaker, which was located centrally with respect to the participant's midline. The computer keyboard was positioned so that the numeric keyboard was in line with the participants' midline. Responses were made using the "6" and "9" keys on the numeric keypad of the keyboard with the right index and middle fingers, respectively. The participant's arm was bent at approximately 90°, with the elbow resting on the table to the right of the keyboard. Because of the tilt of the keyboard with its rear legs extended, the "9" key was above the "6" key (and slightly further from the participant). Half of the participants responded using a spatially compatible S-R assignment, pressing the upper key (the "9" key) when the stimulus started high in pitch (or moved up in pitch) and pressing the lower key (the "6" key) when the stimulus started low (or moved lower). The other half of the participants responded using a spatially incompatible S-R assignment (e.g., pressing the lower key when the stimulus started high in pitch and pressing the upper key when the stimulus started low in pitch).

Procedure. Each participant had two sessions, each of which consisted of a block of 60 practice trials and two blocks of 60 experimental trials. Five repetitions of each of the 12 stimuli were presented in a random order within each block. Accuracy feedback was given on each trial, and overall accuracy was presented at the end of each block. Half of the participants responded to onset pitch for the first session (ignoring the direction of pitch change) and then responded to the direction of pitch change (ignoring onset pitch) in the second session. Half of the participants performed the tasks in the reverse order. During the instruction phase at the beginning of the experiment, the participant heard each of the 12 stimulus sounds once. An accompanying message on the computer screen indicated that "These are the three tones that start high, and become higher in pitch," when the three high and congruent stimuli were played. The same method was used to present the remaining stimuli. Following presentation of the stimuli, we instructed the participant to use one of the two S-R assignments. At the beginning of each session, they were instructed to respond to either onset pitch or direction of pitch change, ignoring the other dimension of the stimulus.

Before the first trial in each block, a "Get Ready!" message was displayed for 1,500 ms. Then, the screen was cleared and the first trial began with the presentation of the auditory stimulus. Immediately following the keypress response, a feedback message was displayed for 1,500 ms, indicating a correct or incorrect response. A 1,500-ms intertrial interval (ITI) preceded presentation of the next stimulus. At the conclusion of the second session, each participant received a brief explanation of the purpose of the study and was dismissed from the experiment.

Results

Practice trials were excluded from the analysis, as were trials on which responses were faster than 100 ms or slower than 2,000 ms (less than 1% of trials). Mean correct RTs and mean accuracy were subjected to an analysis of variance (ANOVA), with order (respond to pitch then to direction or respond to direction then to pitch) and assignment (compatible or incompatible S-R assignment) as between-subjects factors and cue dimension (onset pitch or pitch change), congruity (congruent or incongruent), and separation (small, medium, or large stimulus separation) as within-subjects factors. An alpha level of .05 was used for all statistical tests.

Mean RTs and percentage correct as a function of cue dimension, separation, and congruity are shown in Figures 2 and 3. For the sake of clarity, only significant ($p < .05$) F ratios are presented in the following; occasionally, a nonsignificant result is presented for comparison.

The main effect of cue dimension was significant for both RT and accuracy, reflecting that, overall, responses were faster (676 vs. 832 ms), $F(1, 24) = 13.14, p = .0014, MSE = 155{,}125$, and more accurate (94 vs. 87%), $F(1, 24) = 7.23, p = .0130, MSE = 0.0303$, for onset-pitch than for pitch-change judgments. The main effect of congruity was significant for RT and for accuracy. Mean responses were faster (729 vs. 782 ms), $F(1, 24) = 11.86, p = .0022, MSE = 219{,}503$, and more accurate (94 vs. 88%), $F(1, 24) = 10.89, p = .0031, MSE = 0.0140$, for congruent than for incongruent stimuli.

The Cue Dimension × Congruity interaction did not reach significance for RT, $F(1, 24) = 2.74, p > .11, MSE = 18{,}511$, but it was significant for accuracy, $F(1, 24) = 8.88, p = .0066, MSE = 0.0051$. When attending to onset pitch, responses to congruent stimuli were more accurate than responses to incongruent stimuli (*M*s = 96 vs. 93%, respectively). When attending to pitch change, responses to congruent stimuli were also more accurate than those to incongruent stimuli, and the congruity effect was larger than for onset pitch judgments (*M*s = 93 vs. 83% for congruent and incongruent stimuli, respectively).

As separation increased, responses became significantly faster (*M*s = 793, 744, and 731 ms at the small, medium, and large separations, respectively), $F(2, 48) = 25.06, p = .0001, MSE = 4{,}654$, and more accurate (*M*s = 89, 93, and 91% correct at the small, medium, and large separations, respectively), $F(2, 48) = 15.79, p = .0001, MSE = 0.0026$. The main effect of separation was qualified by significant Cue Dimension × Separation interactions for RT, $F(2, 48) = 44.87, p = .0001, MSE = 3{,}666$, and accuracy, $F(2, 48) = 45.38, p = .0001, MSE = 0.0027$. For RT, the interaction reflects that the pattern of faster responses at greater stimulus separations held only for onset-pitch judgments (*M*s = 760, 651, and 609 ms for small, medium, and large stimulus separations, respectively) and not for pitch-change judgments (*M*s = 827, 831, and 839 ms for small, medium, and large separations, respectively).

Consistent with the RT results, accuracy increased as a function of separation for onset-pitch judgments (*M*s = 88, 98, and 99% for small, medium, and large stimulus separations, respectively) and decreased as a function of separation for pitch-change judgments (*M*s = 90, 88, and 84% for

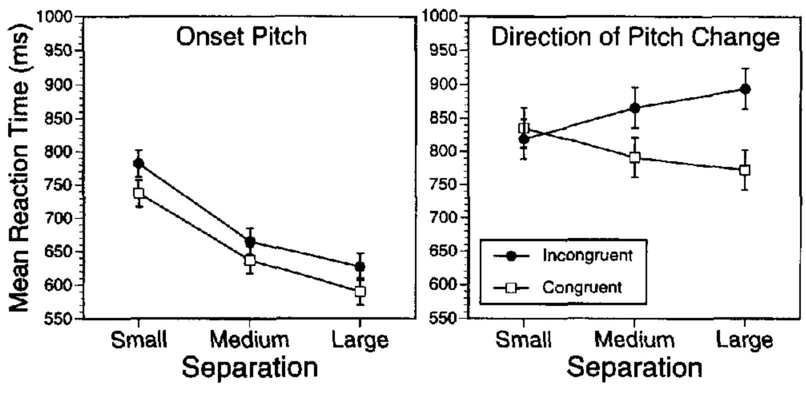

Figure 2. Mean reaction time (in milliseconds) in Experiment 1 as a function of cue dimension, congruity, and separation. Error bars represent standard error of the mean.

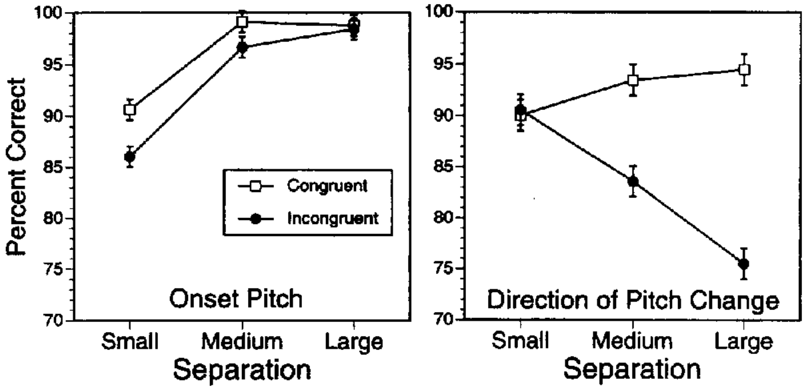

Figure 3. Mean accuracy (percent correct) in Experiment 1 as a function of cue dimension, congruity, and separation. Error bars represent standard error of the mean.

small, medium, and large stimulus separations, respectively). The Congruity × Separation interaction was also significant for RT, $F(2, 48) = 7.98$, $p = .0011$, $MSE = 4{,}194$, and accuracy, $F(2, 48) = 5.43$, $p = .0076$, $MSE = 0.0033$.

The Cue Dimension × Congruity × Separation interaction was significant for both RT, $F(2, 48) = 12.88$, $p = .0001$, $MSE = 3{,}379$, and accuracy, $F(2, 48) = 8.74$, $p = .0007$, $MSE = 0.0052$. As shown in Figure 2 for RT and in Figure 3 for accuracy, when attending to onset pitch (left side of both figures), responses to both congruent and incongruent stimuli were faster (congruent *M*s = 738, 637, and 591 ms; incongruent *M*s = 783, 665, and 627 ms for small, medium, and large separations, respectively) and more accurate (congruent *M*s = 91, 99, and 99%; incongruent *M*s = 86, 97, and 98%) with an increased stimulus separation. However, when attending to pitch change (right side of Figure 2), increased stimulus separation led to faster responding only for congruent stimuli (*M*s = 836, 792, and 772 ms for small, medium, and large separations, respectively), whereas for incongruent stimuli, increased separation led to slower responding (*M*s = 820, 867, and 895 ms for small, medium, and large separations, respectively). The accuracy results mirrored the RT patterns: Accuracy for onset pitch increased with separation for both congruent and incongruent trials, and accuracy for pitch-change responses increased with separation for congruent stimuli (*M*s = 90, 94, and 95%) but decreased with increased separation for incongruent stimuli (*M*s = 91, 84, and 76%; see Figure 3, right side).

The main effect of assignment was not significant for RTs, $F(1, 24) < 1$, nor did assignment appear in any significant interaction. For accuracy, the main effect of assignment was not significant, $F(1, 24) < 1$, but assignment did appear in two significant higher order interactions, the Assignment × Congruity × Separation interaction, $F(2, 48) = 4.04$, $p = .0240$, $MSE = 0.0033$, and the Assignment × Order × Cue Dimension × Separation interaction, $F(2, 48) = 3.47$, $p = .0391$, $MSE = 0.0027$, discussed later. Underlying the Assignment × Congruity × Separation interaction is the finding that responses to incongruent stimuli in the compatible mapping showed a decrease in accuracy with increasing separation (−7% difference between the small and large separations), whereas responses to congruent trials increased in accuracy with increasing separation (9%). In the incompatible mapping, the increase in accuracy with increased separation was small but essentially the same for congruent and incongruent stimuli (4 and 3%, respectively).

There was no main effect of the order of sessions for RT, $F(1, 24) = 2.14$, $p > .15$, $MSE = 413{,}590$, or accuracy, $F(1, 24) < 1$. In addition, there were no significant two-way interactions involving order for either RT or accuracy. The three-way Order × Cue Dimension × Separation interaction was the only interaction to reach significance for RT, $F(2, 48) = 6.96$, $p = .0023$, $MSE = 3{,}666$. Regardless of the order of conditions, onset-pitch responses were faster for larger separations (onset pitch-then-pitch change order *M*s = 659, 571, and 547 ms for small, medium, and large separations, respectively; pitch change-then-onset pitch order *M*s = 862, 731, and 669 ms). For pitch-change responses, responding to pitch change first and then to onset pitch resulted in an increase in RTs with greater separation (*M*s = 835, 867, and 861 ms for small, medium, and large separations, respectively). However, responding first to onset pitch and then to pitch change resulted in little change to pitch-change responses with increased separation (*M*s = 820, 795, and 818 ms for small, medium, and large separations, respectively).

For accuracy, the Assignment × Order × Cue Dimension × Separation interaction, mentioned earlier, was the only significant interaction involving order. In general (i.e., for both compatible and incompatible response assignments), responding to direction of pitch change first and then to onset pitch resulted in greater accuracy for the pitch-change judgments. The increase in accuracy for onset pitch with increased separation was essentially the same for both orders, with both response assignments. The interaction results from the fact that the improvement in accuracy because of responding first to pitch change, as opposed to onset pitch, was larger for the incompatible response assignment groups than for the compatible response assignment. Accuracy with the incompatible assignment, when responding first to direction of pitch change, was essentially unchanged with increased separation (*M*s = 96, 95, and 96% for small, medium, and large separations, respectively), whereas in the other conditions accuracy tended to be lower overall (ranging from 77 to 92%) and tended to decrease somewhat (by 7 to 9%) with increased separation. It may be that participants who responded to direction first, using an incompatible response assignment, recoded the task as "move in the opposite direction as the sound," and this may have affected either their listening or responding strategy, thus making the relative separation information less intrusive.

Discussion

Overall, responses to onset pitch were faster than responses to direction of pitch change. This is, at least in part, likely due to the fact that responding to pitch change required waiting until the pitch changed (i.e., a minimum of 15 ms, the amount of time that the stimulus remained at the onset pitch). However, that does not explain the fact that responses to pitch change were more than 150 ms (i.e., more than the duration of the stimulus) slower on average. It is possible that onset pitch was, overall, more discriminable than direction of pitch change, which could lead to faster responses. Participants may also have had more prior experience making relative pitch judgments than pitch-change judgments, with the result that performance in the onset-pitch task may have benefited from a preexperimental practice effect. Interestingly, Ehrenstein (1994) found very similar results for dynamic visual stimuli. In that experiment, responses to the direction of motion of a displayed dot were on average 100 ms slower than responses to the dot's onset position. Further research may investigate whether this phenomenon is robust and if it is seen for judgments of change in other stimulus dimensions.

The dimensions of pitch and pitch change interacted, such that responses were almost always faster when the direction of pitch change matched the onset pitch (i.e., for congruent stimuli), although this finding was qualified by the Cue Dimension × Separation × Congruity interaction (see Figure 2). The Cue Dimension × Separation × Congruity interaction reflects that onset pitch affected pitch-change judgments more than pitch change affected onset-pitch judgments. In particular, onset-pitch judgments were faster for congruent stimuli than for incongruent stimuli, and for both congruent and incongruent stimuli, the greater the separation, the faster the responses to onset pitch (left panel of Figure 2). When attending to pitch change, a different pattern of results was obtained. Congruent stimuli were, on the whole, still faster than incongruent stimuli, but there was little gain in performance with increased separation (right panel of Figure 2, open squares). In fact, when onset pitch was incongruent with the direction of pitch change, greater separation led to slower responses to pitch change. Hence, it appears that the greater the separation of the pitches from the middle of the pitch space, the more salient relative pitch becomes and the greater its interference with pitch-change judgments. At small separations, where relative pitch is close to the average pitch for the stimulus set, it is apparently easier to ignore the irrelevant onset pitch information, as indicated by a lack of a congruity effect.

As mentioned, we chose to use 12 stimuli to prevent listeners from learning responses to specific stimuli. Supporting the assertion that the congruity and separation of the stimuli are the relevant dimensions affecting performance in this task and not peculiarities of specific stimuli, analyses performed on the stimuli examining the dimensions of relative pitch (high vs. low), direction of pitch change (higher vs. lower), and separation (small, medium, or large, as defined above) showed that there were no significant interactions involving specific stimuli, and only separation had a significant main effect on RT.

The results of this experiment clearly confirmed our predictions of interactions between the dimensions of onset pitch and pitch change in dynamic auditory stimuli. That is, selective attention to either onset pitch or direction of pitch change is not perfect. The uniformity of the effect of congruity on onset-pitch judgments further suggests that it is the change in perceived pitch (which was equal at each separation) and not the change in absolute frequency (which varied from 185 to 927 Hz) that is processed when judging auditory stimuli.

The lack of S-R assignment effects in the present experiment is surprising. If pitch were indeed treated by listeners as having a spatial aspect (e.g., Mudd, 1963; R. Walker, 1987), then there should have been spatial compatibility effects resulting from the different assignments of stimuli to responses. If there is in fact no effect of assigning upper keypress responses to high-pitched rather than low-pitched sounds, and vice versa for lower keypresses, that would suggest that pitch is not spatially coded, or at least that the pitch space is not directly comparable with the response space in this case. Another possibility is that it was the responses that were not treated as having a vertical (upward and downward response) aspect. This possibility is examined in the following experiments.

Experiment 2

Experiment 2 was conducted to investigate the lack of S-R compatibility effects in Experiment 1. S-R compatibility effects were expected to arise from the assignment of high and low (or increasing and decreasing) pitches to vertically oriented keypress responses. The lack of such effects suggests that the stimuli, the responses, or both were not treated by the participants as having a vertically oriented spatial aspect. Recent evidence (Lippa, 1996) suggests that the responses used in Experiment 1 may have been coded as right and left rather than as up and down. Lippa (1996) showed that when the hand of the participant is positioned such that the allowable responses are to the right and left relative to the responding hand (as was the case in Experiment 1, in which the arm was bent to put the index finger on the "6" key and the second finger on the "9" key of the numeric keypad), responses are coded as if they were aligned horizontally rather than vertically. Because the dimensional overlap between vertical stimuli and horizontal responses is relatively low, S-R compatibility effects are often reduced or absent in such situations (Lippa). In Experiment 2, we changed the response set, making it clearly vertical in a physical sense, relative to both the responding hand and the whole body. Such a vertical response set should correspond more closely to the pitch space, if pitch is represented as occupying a vertically oriented space, thus providing the basis for spatial S-R compatibility effects.

Method

Experiment 2 used the same stimulus set as Experiment 1 but a different response apparatus. Instead of using the keyboard to make responses, participants responded by pressing buttons on a vertical response panel (see Figure 4). In an attempt to manipulate the salience of upward and downward movements, two different sets of response buttons, differing in their separation relative to each other, were used. As in Experiment 1, response assignment (compatible or incompatible) was manipulated between-subjects. Because of a suggestion of differential carryover effects in Experiment 1, cue dimension was changed to a between-subjects variable.

Participants. There were 88 new participants from the same participant pool used in Experiment 1.

Apparatus. The computer apparatus that controlled the experiment and presented the sound stimuli was identical to that used in Experiment 1, except that participants now responded by pressing the buttons of a five-button Serial Response Box (a response device produced by Psychology Software Tools, Pittsburgh, PA, specifically for use with the MEL software), which was supported on a wooden stand as illustrated in Figure 4. The response buttons were arranged in a single vertically oriented row, in approximately the same plane as the front of the computer screen (although the whole response box was inclined away from the participant at a slight angle to allow the participant to press the buttons with the fleshy part of the fingertip rather than the tip of the fingernail). For the purposes of explanation, the buttons are referred to here in numerical order, starting from the bottom. Thus, the lowest button is Button 1 and the top button is Button 5 (they were not so labeled for the participants).

Participants rested their right wrist on a 5-cm thick piece of foam placed on the table between the participant and the response box. This maintained a comfortable arm and hand position while allowing for unimpeded movement to any of the response buttons. Specifically, the hand was oriented palm-down, with the thumb and last three fingers held in a loose fist. The index finger was held straight out (i.e., pointing) and perpendicular to the planes of the body and the response buttons (which were essentially parallel). The right elbow was held as close to the body as was comfortable, so that the forearm was very nearly collinear with the index finger, with the wrist only very slightly deviated, if at all. All responses were made with the tip of the right index finger. Two response sets were used: "near" responses, which involved moving from Button 3 (the "home" button) to either Button 2 or Button 4 (see Figure 4), and "far" responses, which involved moving from the home button to either Button 1 or Button 5.

Stimuli. The stimuli were identical to those used in Experiment 1.

Procedure. The trial procedure differed from Experiment 1 only as follows: At the start of each trial, the light to the side of Button 3 of the response box came on (see Panel A of Figure 4), and a message instructed the participant to press and hold the home button. When the home button was pressed and held, the middle light was turned off and the lights were turned on beside Buttons 2 and 4 (in the near-response condition) or beside Buttons 1 and 5 (in the far-response condition) as a reminder of the valid responses for that trial (Panel B of Figure 4), while a "Get Ready!" message was displayed on the computer screen for 1,000 ms. The stimulus was presented, and, as quickly as possible, the index finger was to be moved off the home button and to one of the valid response buttons. If the participant pressed the correct button, the light beside it stayed on while the other lights were turned off, and the message, "Correct response" appeared on the computer screen (see Panel C of Figure 4). If the subject pressed an incorrect button (or made no response within 3 s), all the response box lights came on, and "Wrong response!" appeared on the computer screen. Finally, if the participant released the home button before the stimulus started, the message, "You moved too soon!" appeared on the computer screen. In all cases, the feedback message was displayed for 1,500 ms, followed by a 1,500-ms ITI.

The block structure was the same as in Experiment 1, with three blocks of 60 trials in each condition of the experiment. Instead of responding to a different cue dimension in each session, as was the case in Experiment 1, participants now responded to the same cue dimension, either onset pitch or pitch change, for the entire experiment.

As in Experiment 1, half the participants responded throughout with a compatible response assignment, whereas the other half responded throughout with an incompatible assignment. Within each assignment condition, half the participants responded using the near-response buttons and half used the far-response buttons.

Results

The practice trials and trials on which responses were less than 100 ms or greater than 3,000 ms (less than 1% of trials) were excluded from the analyses. Mean correct RTs and mean accuracy were subjected to separate ANOVAs, with cue dimension (onset pitch vs. pitch change), assignment (compatible vs. incompatible), and response type (near vs. far) as between-subjects factors. Congruity (congruent or incongruent) and separation (small, medium, or large) were within-subjects factors. Total RT (the time from the onset of

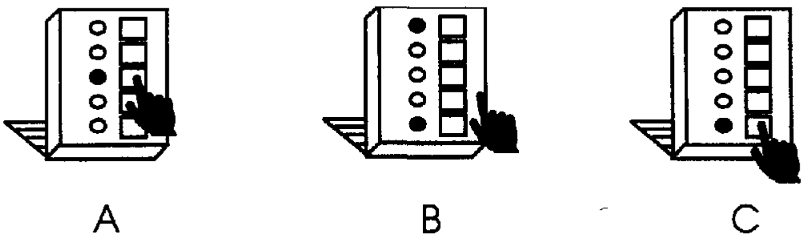

Figure 4. Schematic diagram of the apparatus used in Experiment 2.

the sound until the response button press) and accuracy were the primary dependent measures, as before. The total time to respond is of most relevance in practical applications. However, in addition to total RT, the time required to release the home button (lift time), and the time required to move to the response button (movement time) can be of theoretical interest, especially because lift time is where one may expect compatibility effects to arise. Thus, these two additional variables were also analyzed. It is important to note that to avoid confusion with response time, especially when abbreviated, the term *lift time* is used here to refer to what is often called *reaction time* in the literature (e.g., Fitts & Peterson, 1964).

Mean correct RTs, as a function of cue dimension, separation, and congruity, are shown in Figure 5. The main effect of cue dimension was significant, reflecting that overall, responses were faster for onset-pitch judgments than for pitch-change judgments (617 vs. 820 ms, respectively), *F*(1, 80) = 25.26, *p* = .0001, *MSE* = 215,811. The main effect of cue dimension was not significant for accuracy, *F*(1, 80) = 3.02, *p* = .0862, *MSE* = 0.0430.

The significant main effects of congruity for both RT, *F*(1, 80) = 53.36, *p* = .0001, *MSE* = 9,404, and accuracy, *F*(1, 80) = 28.62, *p* = .0001, *MSE* = 0.0177, reflect that overall, responses to congruent stimuli were faster (*M*s = 688 vs. 749 ms) and more accurate (*M*s = 96 vs. 90%) than responses to incongruent stimuli.

The Cue Dimension × Congruity interaction also was significant for both RT, *F*(1, 80) = 12.15, *p* = .0008, *MSE* = 9,404, and accuracy, *F*(1, 80) = 5.86, *p* = .0178, *MSE* = 0.0177. When attending to onset pitch, responses to congruent stimuli were 32 ms faster than responses to incongruent stimuli (*M*s = 601 vs. 633 ms), and accuracy was higher for congruent than for incongruent trials (*M*s = 96 vs. 93%, respectively). When attending to pitch change, however, the difference between RTs to congruent and incongruent stimuli was 91 ms (*M*s = 775 vs. 866 ms for congruent and incongruent, respectively), and the difference in accuracy was also larger (*M*s = 96 and 87%, for congruent and incongruent trials, respectively).

The main effect of separation, *F*(2, 160) = 13.52, *p* = .0001, *MSE* = 4,022, reflects that RTs were shorter at the larger separations (*M*s = 739, 706, and 711 ms at the small, medium, and large separations, respectively). Consistent with this, accuracy was higher at the larger separations (*M*s = 90, 95, and 93% for small, medium, and large separations, respectively), *F*(2, 160) = 24.33, *p* = .0001, *MSE* = 0.0034.

The Cue Dimension × Separation interaction was significant for RT, *F*(2, 160) = 55.06, *p* = .0001, *MSE* = 4,022. When onset pitch was the relevant cue dimension, RTs decreased with larger separation (*M*s = 675, 599, and 577 ms for small, medium, and large separations, respectively). However, when attending to pitch change, RTs increased with larger separation (*M*s = 802, 813, and 845 ms for small, medium, and large separations, respectively). The Cue Dimension × Separation interaction was also significant for accuracy, *F*(2, 160) = 86.32, *p* = .0001, *MSE* = 0.0034, with a similar pattern of results. When attending to onset pitch, accuracy increased with larger separation (*M*s = 87, 97, and 98% correct for small, medium, and large separations, respectively). When pitch change was the relevant cue dimension accuracy decreased with greater separation (*M*s = 93, 92, and 88% for small, medium, and large separations, respectively).

The Congruity × Separation interaction was significant for both RT, *F*(2, 160) = 14.22, *p* = .0001, *MSE* = 4,408, and accuracy, *F*(2, 160) = 5.86, *p* = .0035, *MSE* = 0.0056. For congruent stimuli, performance was better with increasing separation (RT *M*s = 725, 678, and 660 ms; accuracy *M*s = 92, 97, and 98% for small, medium, and large separations, respectively). However, for incongruent stimuli, no consistent patterns were evident (RT *M*s = 752, 734, and 762 ms; accuracy *M*s = 88, 92, and 89% at the small, medium, and large separations, respectively). The Cue Dimension × Congruity × Separation interactions for RT, *F*(2, 160) = 24.84, *p* = .0001, *MSE* = 4,408, and accuracy, *F*(2, 160) = 24.61, *p* = .0001, *MSE* = 0.0056, reflect that there was a Congruity × Separation interaction only for responses to pitch change (see Figure 5). Overall, the patterns of results for the three-way Cue Dimension × Congruity × Separation interactions for RT and accuracy follow those found in Experiment 1.

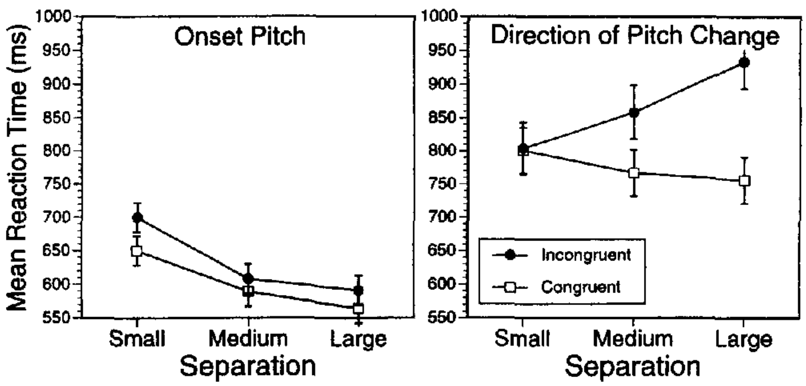

Figure 5. Mean reaction time (in milliseconds) in Experiment 2 as a function of cue dimension, congruity, and separation. Error bars represent standard error of the mean.

The major difference in results between this experiment and Experiment 1 is that here the main effect of assignment was significant for RT, $F(1, 80) = 5.43, p = .0224, MSE = 215{,}811$, with compatible responses faster than incompatible responses (*M*s = 671 vs. 766 ms, respectively). Assignment did not have a significant main effect on accuracy, $F(1, 80) < 1$. The only significant interaction with assignment for RT or accuracy was the three-way Assignment × Cue Dimension × Separation interaction for accuracy, $F(2, 160) = 4.42, p = .0136, MSE = .0034$. For responses to pitch, accuracy was lowest at the smallest separation for both assignments (compatible accuracy *M*s = 86, 98, and 99%; incompatible accuracy *M*s = 89, 97, and 98% for small, medium, and large separations, respectively). The corresponding changes in accuracy for pitch-change judgments were as follows: compatible *M*s = 93, 91, and 87%, and incompatible *M*s = 93, 93, and 90%.

The significant main effect of response type for RT reflects that, overall, responding to the near buttons was faster than responding to the far buttons (664 vs. 773 ms), $F(1, 80) = 7.24, p = .0087, MSE = 215{,}811$. The main effect of response type was not significant for accuracy, $F(1, 80) < 1$, nor were there any significant interactions involving response type for accuracy. The only interaction with response type to reach significance was the Cue Dimension × Separation × Response Type interaction for RT, $F(2, 160) = 3.31, p = .0388, MSE = 4{,}022$, which reflects that separation generally had a larger effect on RT with the far responses than with the near responses. For responses to onset pitch, there was a 110-ms speedup going from small to large separations for far responses and an 80-ms speedup from small to large separations for near responses. When attending to pitch change, there was a 30-ms slowing between small and large separations for near responses and a 60-ms slowing for far responses.

The results of the analyses of lift time (LT) and movement time (MT) were almost identical to the RT results, therefore only departures are reported here. First, for LT, the main effects of response type, $F(1, 80) = 2.21, p = .14, MSE = 109{,}860$, and assignment, $F(1, 80) = 1.48, p = .23, MSE = 109{,}860$, were not significant. We note, however, that LTs with an incompatible assignment were numerically faster than with a compatible assignment (350 vs. 385 ms, respectively). The Assignment × Congruity interaction was significant for LT, $F(1, 80) = 9.73, p = .0025, MSE = 1{,}480$. With a compatible response assignment, LTs to congruent stimuli averaged 26 ms faster than to incongruent stimuli (372 vs. 398 ms). With an incompatible assignment, LTs to congruent stimuli averaged only 4 ms faster than to incongruent stimuli (348 vs. 352 ms).

For MT, the only difference of note from the RT analysis was the significant Assignment × Cue Dimension interaction, $F(1, 80) = 5.66, p = .0198, MSE = 134{,}430$. When attending to onset pitch, the movement part of incompatible responses averaged 53 ms slower than the movement part of compatible responses (316 vs. 263 ms for incompatible and compatible responses, respectively). However, when attending to direction of pitch change, MTs for incompatible responses averaged 205 ms slower than for compatible responses (515 vs. 310 ms, respectively).

Discussion

The patterns of results for RT and accuracy, with respect to cue dimension, congruity, and separation, were identical to Experiment 1. This is perhaps not surprising, because the same stimuli were used in both experiments. However, it is important to note that participants in Experiment 2 responded only to one or the other of the cue dimensions. The fact that this change had no effect on the overall pattern of results lends credibility to the hypothesis that congruity effects with the dimensions of pitch and pitch change are robust.

The significant main effect of assignment in this experiment shows that spatial compatibility effects can arise from the interaction of pitch and a vertically oriented response set. The presence of such a compatibility effect in Experiment 2, but not in Experiment 1, and the fact that the only relevant change from Experiment 1 to 2 was to make the response buttons and finger movement vertical in the frontal plane, provides evidence that pitch is indeed perceived as having a vertically aligned spatial aspect.

An interesting detail of the results is that the effects of the irrelevant stimulus dimension depended to some extent on response assignment. In particular, the Assignment × Cue Dimension × Separation interaction, such that separation in onset pitch had a bigger effect in the compatible assignment than in the incompatible assignment, suggests that requiring participants to use an incompatible response assignment also seems to suppress the processing of response-congruent, but irrelevant, information. The Assignment × Congruity interaction for LT, such that LTs showed a congruity effect only with the compatible assignment, is also consistent with this supposition and with the general finding that the effects of the correspondence of an irrelevant stimulus dimension to responses is often reduced or reversed when the S-R assignment is incompatible (De Jong, Liang, & Lauber, 1994; Simon et al., 1976). It should be noted that the listeners in our experiment were not told how to structure their speeded responses in any way. That is, some may have decided on the correct response before initiating a ballistic movement, whereas others may have lifted their finger from the home key and decided in the air which response button to press. Indeed, it may have varied from trial to trial for some participants. This makes the specific interpretation of LT and MT less meaningful; thus, we continue to consider the total RT as the primary chronometric in this report.

One final comment concerns the curious result that LTs with the incompatible assignment were numerically faster than with the compatible assignment. Even though this result was not statistically significant, as already reported, it is still a fairly large effect (35 ms) in what one might consider the wrong direction (i.e., inconsistent with the other results reported here). We are not sure why the between-groups manipulation of assignment showed such an effect on LT, but by the time the response was completed (i.e., looking at the RT data), the results were as anticipated.

Experiment 3

Although it seems likely that the only substantive difference between Experiments 1 and 2 was that the alignment of the response buttons was not truly vertical in the first experiment whereas it was in the second experiment, it is also the case that there was a difference in the response action itself. In particular, the response in Experiment 1 involved moving one of two fingers, with the responding hand remaining essentially static. The response in Experiment 2, however, required moving the whole hand, and responding always with the index finger. Thus, it is possible that an element of the responses other than the spatial layout of the buttons was the cause of the differences in spatial compatibility effects seen in Experiments 1 and 2. To investigate this possibility, we replicated Experiment 2 with identical response actions but with the response buttons aligned horizontally rather than vertically.

Method

Experiment 3 included 68 new participants from the same participant pool and was identical to Experiment 2 except for the following changes. The response box was rotated clockwise 90° in the same plane as the front of the computer screen, so that the buttons were still facing the participant but were now arranged left to right rather than bottom to top. As before, half the participants responded to the onset pitch of the stimulus, and half responded to the direction of pitch change. Now, within each of these groups, half of the listeners responded to a high (or becoming higher) pitch with a right button press and responded to a low (or becoming lower) pitch with a left button press. The other half of the participants responded using the opposite S-R assignment (i.e., high to left, low to right). One final change in this experiment was that participants only responded with the near responses. Thus, response type was not a factor in the design.

Results

As in the analysis of the previous experiment, the practice trials and trials on which responses were less than 100 ms or greater than 3,000 ms (less than 1% of trials) were excluded. Mean correct RTs and mean accuracy were subjected to separate ANOVAs, with cue dimension (onset pitch vs. pitch change) and assignment (high to right vs. high to left) as between-subjects factors. Stimulus congruity (congruent or incongruent) and separation (small, medium, or large) were within-subjects factors.

The pattern of both RT and accuracy results for cue dimension, separation, and congruity were essentially the same as those in Experiments 1 and 2. For the sake of comparison, RT results are shown in Figure 6 as a function of cue dimension, separation, and congruity. Again, responses were faster for onset-pitch judgments than for pitch-change judgments (625 vs. 717 ms, respectively), $F(1, 64) = 4.06$, $p = .0481$, $MSE = 214{,}316$. As before, there was a nonsignificant trend toward more accurate responses for onset-pitch judgments than for pitch-change judgments (93% vs. 90%, respectively), $F(1, 64) = 2.72$, $p = .1042$, $MSE = 0.0367$. The significant main effects of congruity for both RT, $F(1, 64) = 31.62$, $p = .0001$, $MSE = 16{,}145$, and accuracy, $F(1, 64) = 18.91$, $p = .0001$, $MSE = 0.0231$, reflect that overall, responses to congruent stimuli were faster (*M*s = 636 vs. 706 ms) and more accurate (*M*s = 95 vs. 88%) than responses to incongruent stimuli.

Although the data did show the same pattern as in Experiments 1 and 2, the Cue Dimension × Congruity interaction in this experiment did not reach significance for either RT, $F(1, 64) = 2.94$, $p = .0914$, $MSE = 16{,}145$, or accuracy, $F(1, 64) = 2.38$, $p = .1276$, $MSE = 0.0231$. It is important to note, however, that the Cue Dimension × Congruity × Separation interactions, discussed later, were significant.

The main effect of separation, $F(2, 128) = 21.93$, $p = .0001$, $MSE = 2{,}859$, shows that RTs were shorter at the larger separations (*M*s = 695, 661, and 656 ms at the small, medium, and large separations, respectively). Accuracy also tended to be higher at the larger separations (*M*s = 89, 94, and 92% for small, medium, and large separations, respectively), $F(2, 128) = 15.07$, $p = .0001$, $MSE = 0.0051$.

The Cue Dimension × Separation interaction was significant for RT, $F(2, 128) = 56.90$, $p = .0001$, $MSE = 2{,}859$, and for accuracy, $F(2, 128) = 43.83$, $p = .0001$, $MSE = 0.0051$, with identical patterns of results as in Experiments 1 and 2. The Congruity × Separation interaction did not reach

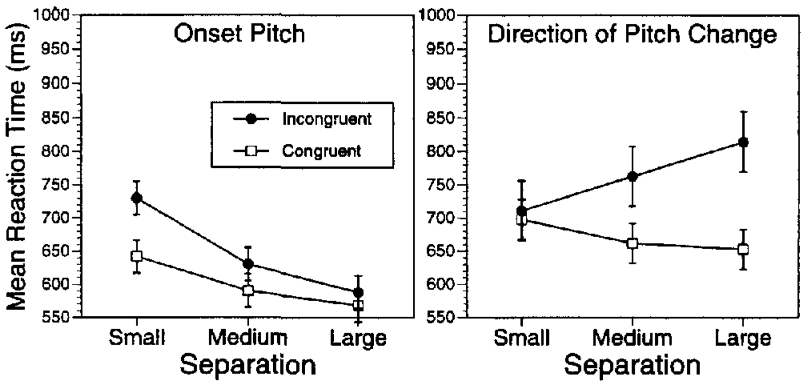

Figure 6. Mean reaction time (in milliseconds) in Experiment 3 as a function of cue dimension, congruity, and separation. Error bars represent standard error of the mean.

significance for either RT, $F(2, 128) = 2.63$, $p = .0759$, $MSE = 5{,}356$, or accuracy, $F(2, 128) = 2.48$, $p = .0879$, $MSE = 0.0060$. The Cue Dimension × Congruity × Separation interaction was significant for both RT, $F(2, 128) = 19.07$, $p = .0001$, $MSE = 5{,}356$, and accuracy, $F(2, 128) = 28.53$, $p = .0001$, $MSE = 0.0060$, with the overall patterns of results for this interaction the same as in both Experiments 1 and 2.

One should recall that the only major procedural difference between this experiment and Experiment 2 was in the alignment of the response buttons. In the present experiment, there was no difference between the two S-R assignments for RT (*M*s = 667 and 675 ms for high to right and high to left, respectively), $F(1, 64) < 1$, or accuracy (*M*s = 91 and 92%, respectively), $F(1, 64) < 1$. The only significant interaction involving assignment was the Assignment × Separation interaction for RT, $F(2, 128) = 3.51$, $p = .0328$, $MSE = 2{,}859$. In both assignments, RTs decreased with increased separations, but the overall improvement was larger in the high to right assignment (high to right *M*s = 702, 654, and 646 ms for small, medium, and large separations, respectively; high to left *M*s = 689, 668, and 665 ms, respectively). The Assignment × Separation interaction was not significant for accuracy, $F(2, 128) < 1$, and no other interactions with assignment reached significance for RT or accuracy.

Discussion

The stimuli in this experiment were identical to those used in Experiments 1 and 2. The general pattern of results relating to stimulus dimension interactions, namely the effects of cue dimension, congruity, and separation, closely followed those of Experiments 1 and 2. More relevant to the current purpose is that the effect of assignment found in Experiment 2 was not replicated. Rather, as in Experiment 1, there was no evidence of spatial compatibility effects for dynamic pitches and horizontal responses. Thus, it seems that compatibility effects with pitch stimuli are only likely to arise when the response set has a vertical aspect.

General Discussion

Whether listeners can attend to relevant aspects of an auditory stimulus while ignoring irrelevant aspects is of prime importance in the design of the auditory displays that are used in the interpretation and analysis of many types of data. In particular, because pitch is the most commonly used dimension in these displays, it is crucial to understand how well listeners can selectively attend to pitch and to changes in pitch while ignoring other changes in the display sounds. Furthermore, if auditory displays are to represent dynamic data successfully, it is critical to know how changes in displayed information (and therefore, changes in the display pitches) are perceived and how they interact with other aspects of the display. Finally, many tasks involving auditory displays also include manual control actions in response to the stimuli. In such cases, interface designers must be able to understand how the perception of stimulus dimensions and the interactions between stimulus dimensions affect the physical responses made by the listener.

The three experiments reported here examined the manner in which the auditory dimensions of pitch and pitch change interact to influence performance in a selective listening task. In each experiment, listeners were instructed to attend either to the relative pitch (i.e., whether the stimulus started high or low in pitch) or to the direction of pitch change (i.e., whether the stimulus became higher or lower in pitch) of dynamic stimuli and then to respond both quickly and accurately to the attended cue dimension.

In all three experiments, attributes of the sound stimuli interacted in their influence on the speed and accuracy of responses. Overall, responses were faster for judgments of pitch than for judgments of pitch change. However, the most important finding was that relative pitch and pitch change interacted to produce congruity effects. Responses were relatively slow when the to-be-ignored stimulus dimension conflicted with the to-be-attended stimulus dimension (i.e., for incongruent stimuli), which indicates that participants were unable to listen to either aspect of the sounds in a completely selective manner. Moreover, the intrusion of the information in one dimension onto judgments regarding the other dimension was not symmetrical. Pitch information had a greater influence on responses to direction of pitch change than pitch-change information had on judgments of relative pitch.

Another important factor was the relative separation of a given sound from the middle pitch of the stimulus set. The magnitude of the congruity effect was fairly uniform across the range of stimulus separations for onset-pitch judgments. However, for pitch-change judgments, responses to the stimuli that had greater separations were more influenced by irrelevant pitch information than responses to small-separation stimuli. This effect of separation on the congruity effect only occurred for judgments of pitch change, which suggests that onset pitch may have been more salient than pitch change. However, it should also be noted that in the stimulus set used in these experiments, the amount of pitch change was not varied across stimuli whereas the relative separation of the onset pitches was.

The factors of congruity, separation, and cue dimension had nearly identical effects in all three experiments. However, the way in which stimuli were assigned to responses had different effects. In Experiment 1, in which listeners responded to the stimuli using the "6" and "9" keys on the numeric keypad of a computer keyboard, there was no effect of whether the upper key (the "9"), for example, was mapped to the high-pitch stimuli and the lower key (the "6") was mapped to the low-pitch stimuli, compared with the inverse S-R mapping. This lack of an assignment effect in Experiment 1 was initially somewhat surprising, given the canonical finding of S-R compatibility effects between vertically arranged stimuli and vertically arranged responses (e.g., Proctor & Reeve, 1990). We had anticipated effects of the assignment of stimuli to responses, under the hypothesis that higher pitches would be treated by listeners as corresponding to higher spatial positions and lower pitches treated as corresponding to lower spatial positions. The fact that such spatial S-R compatibility effects were obtained in Experiment 2, in which a vertical set of buttons was used to respond, indicates that these stimuli were indeed being treated as having a spatial aspect. Clearly, then, the specific choice of responses can affect performance in such tasks involving dynamic sound stimuli.

It seems likely that in Experiment 1, the responses were actually coded as left and right responses rather than up and down responses because the participant's hand and arm position resulted in the response keys being aligned left–right relative to the hand, despite the up–down alignment of the keys relative to the whole body. This interpretation is in keeping with Lippa's (1996) conclusion that spatial coding of responses is made relative to the participant's hand and not relative to the sagittal plane. The lack of spatial S-R compatibility effects in Experiment 1 would therefore be attributable to the vertical stimuli being assigned to effectively horizontal responses. Even if the listeners in Experiment 1 treated the onset pitch or direction of pitch change of the stimuli as being mapped to left and right responses, we might still expect an assignment effect because there have been findings of correspondences between high pitches and right responses and low pitches and left responses—although these effects are much weaker than up–down correspondence effects (e.g., Mudd, 1963; R. Walker, 1987). The complete absence of spatial S-R compatibility effects in Experiment 1 is therefore still somewhat surprising. On the other hand, the finding of assignment effects in Experiment 2, in which responses were vertical relative to the hand (and to the rest of the body), supports the hypothesis that listeners actually do code pitch in a vertical manner, with high pitch corresponding to high spatial position and low pitch corresponding to low spatial position.

In Experiment 3, we used an unambiguously horizontal response set that was more directly comparable with the response set in Experiment 2. If pitch is treated in a vertical manner, then we would anticipate little or no S-R compatibility effects in Experiment 3, in direct contrast to the strong compatibility effects found in Experiment 2. As expected, the pitch stimuli did not interact with the horizontal responses. Taken together, these experiments demonstrate that dynamic pitch stimuli are treated in a spatial manner and are susceptible to the same sorts of S-R compatibility effects seen with other spatial stimuli.

Many types of tasks use auditory displays for the analysis and interpretation of various sorts of data sets. Tasks in which changes in complex data must be studied seem particularly suited to auditory analysis. Indeed, the relative ease with which we understand speech is testimonial to the efficacy of the auditory system as a pattern-recognition device. However, the successful design of auditory displays requires a clearer understanding of how the various attributes of sound stimuli used in these tasks interact to influence perception. In particular, it is crucial to determine how the perception of auditory displays depends on dimensional interactions in dynamic sound stimuli. It should be noted that our experiments used speeded responses, whereas many tasks that are particularly well suited to sonification or auditory displays may not have such time pressure. However, the speeded-classification paradigm is useful in determining underlying population stereotypes regarding the dimensions of the stimuli. It is the knowledge and consideration of those stereotypes that is most crucial in the design of successful auditory displays, regardless of the task requirements.

Previous work (e.g., Melara & O'Brien, 1987; Melara & Marks, 1990b) has shown that several dimensions of auditory stimuli can and do interact in selective-listening tasks. Neuhoff and McBeath (1996) have extended this research into the realm of dynamic auditory stimuli. The findings we report here contribute to the general understanding of auditory processing by extending the list of dimensional interactions to include interactions between pitch and pitch change and, especially, by providing information about the nature of the interactions that arise. In addition, they contribute to an important new literature that should be used to establish practical guidelines for increasingly successful auditory displays.

In Experiments 1, 2, and 3, the performance level was relatively high when pitch was the relevant dimension (e.g., mean accuracies of 94, 94, and 93%, respectively). Although slightly lower than for onset pitch, performance with pitch change was also very good (e.g., mean accuracies of 87, 91, and 90%). This provides experimental validation that pitch is well suited to the display of dynamic data because both pitch and pitch change can be attended to, albeit not with perfect independence. This experimental result is presumably a better reason for selecting pitch as the display dimension than the fact that it happens to be the easiest to control with current hardware. The present work also yields guidance about how to take the interactions of pitch and pitch change into account when designing an auditory display.

An example of producing auditory graphs may help demonstrate how the results of the present experiments can be beneficial in display design. In both visual and auditory displays, the ultimate use for the output must be considered when deciding on the range of values to use in the display dimension. The visual analogy would be the choice of axis minimum and maximum values. A display user might need to determine the general values of the data and ignore relatively minor fluctuations in the value. For example, a radiologist might initially need to determine the current size of a tumor, ignoring for the moment any changes (see, for example, Martins & Rangayyan, 1997; Martins, Rangayyan, Portela, Amaro, & Ruschioni, 1996). If the data values for size are mapped onto pitch, then the task is to attend selectively to relative pitch and ignore pitch changes. That is, "how big is the tumor?" translates to "how high is the pitch?" On the basis of the present findings, we would expect that there may be some intrusion of pitch-change information. Selective attention is not perfect. However, a greater separation in the onset of pitches means better performance on the task, with no increased interference from pitch-change information as a result. In contrast, when the radiologist is later determining the effect of a treatment on tumor size over several weeks, then pitch change is the important dimension. In that case, in terms of performance on a selective-listening task, there is little to be gained with increased pitch separation in the case of congruent stimuli (e.g., a large tumor that is getting larger) but much to be lost in the case of incongruent stimuli (e.g., a large tumor that is getting smaller). Hence, to reduce the deleterious effects of irrelevant pitch information, an auditory display designer might restrict the range over which stimulus onsets may vary for this type of task.

Naturally, there are still many stimulus dimensions that have yet to be investigated before a general theory of interactions in dynamic auditory stimuli can be developed. For example, we have yet to determine how a change in loudness or spatial position affects judgments of the relative loudness or starting spatial position of a sound. In addition, certainly much more research needs to be done on how the findings with auditory displays in one field, or for one type of task, transfer to other domains, or different kinds of data. It should be emphasized that we are still a long way from a "cookbook" of data-to-display mappings for auditory displays.

The finding of S-R compatibility effects in Experiment 2 emphasizes that workstation developers must carefully consider the relationship of controls to auditory display elements as well as to visual ones. Dynamic auditory stimuli, critical in the display of dynamic data, can be treated by listeners as spatial stimuli; hence, designers need to consider the spatial relationships between controls and both visual and auditory display elements. If S-R compatibility would be beneficial to a task involving the auditory display of information, then the results of Experiment 2 demonstrate that such effects can be created. However, the lack of assignment effects in Experiments 1 and 3 suggests that response sets that minimize S-R compatibility effects with dynamic pitch stimuli can also be found. Listeners seem unable to attend to one dimension of a dynamic auditory stimulus in an entirely selective manner, but the present research begins to highlight ways to maximize performance by controlling attributes of both the stimulus and the response and their relation to one another.

References

Albers, M. C. (1996). Auditory cues for browsing, surfing, and navigating the WWW: The audible web. In S. Frysinger & G. Kramer (Eds.), *Proceedings of the third International Conference on Auditory Display, ICAD '96* (pp. 85–90). Palo Alto, CA: ICAD.

Axen, U., & Choi, I. (1995). Using additive sound synthesis to analyze simplicial complexes. In G. Kramer & S. Smith (Eds.), *Proceedings of the second International Conference on Auditory Display, ICAD '94* (pp. 31–43). Santa Fe, NM: ICAD.

Begault, D. R., Wenzel, E. M., Shrum, R., & Miller, J. (1996). A virtual audio guidance and alert system for commercial aircraft operations. In S. Frysinger & G. Kramer (Eds.), *Proceedings of the third International Conference on Auditory Display, ICAD '96* (pp. 117–122). Palo Alto, CA: ICAD.

Blattner, M., Greenberg, R., & Kamegai, M. (1990). Listening to turbulence: An example of scientific audiolization. In *Proceedings of the ACM SIGCHI 90 workshop on multimedia and multimodal interface design* (pp. 1–8). New York: Association for Computing Machinery.

Clark, H. H., & Brownell, H. H. (1976). Position, direction, and their perceptual integrality. *Perception & Psychophysics, 19,* 328–334.

Deutsch, D. (Ed.). (1982). *The psychology of music.* New York: Academic Press.

De Jong, R., Liang, C.-C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus–response correspondence. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 731–750.

Dyer, F. N. (1972). Latencies for movement naming with congruent and incongruent word stimuli. *Perception & Psychophysics, 11,* 377–380.

Ehrenstein, W. H. (1994). The Simon effect and visual motion. *Psychological Research, 56,* 163–169.

Fitts, P. M., & Deininger, R. L. (1954). S-R compatibility: Correspondence among paired elements within stimulus and response codes. *Journal of Experimental Psychology, 48,* 483–491.

Fitts, P. M., & Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology, 67,* 103–112.

Flowers, J. H., Buhman, D. C., & Turnage, K. D. (1997). Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. *Human Factors, 39,* 341–351.

Flowers, J. H., & Hauer, T. A. (1992). The ear's versus the eye's potential to assess characteristics of numeric data. Are we too visuocentric? *Behavior Research Methods, Instruments, & Computers, 24,* 258–264.

Flowers, J. H., & Hauer, T. A. (1993). "Sound" alternatives to visual graphics for exploratory data analysis. *Behavior Research Methods, Instruments, & Computers, 25,* 242–249.

Flowers, J. H., & Hauer, T. A. (1995). Musical versus visual graphs: Cross-modal equivalence in perception of time series data. *Human Factors, 37,* 553–569.

Forbes, T. W. (1946). Auditory signals for instrument flying. *Journal of the Aeronautical Society, May,* 255–258.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1,* 225–241.

Heister, G., Schroeder-Heister, P., & Ehrenstein, W. (1990). Spatial coding and spatio-anatomical mapping: Evidence for a hierarchical model of spatial stimulus–response compatibility. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus–response compatibility: An integrated perspective* (pp. 117–143). Amsterdam: North Holland.

Hellier, E. J., Edworthy, J., & Dennis, I. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human Factors, 35,* 693–706.

Hommel, B. (1994). Spontaneous decay of response-code activation. *Psychological Research, 4,* 261–268.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: A cognitive basis for a model and taxonomy of stimulus–response compatibility. *Psychological Review, 97,* 253–270.

Kramer, G. (1994). Some organizing principles for representing data with sound. In G. Kramer (Ed.), *Sonification, audification, and auditory interfaces* (pp. 185–221). Reading, MA: Addison Wesley.

Leung, Y. K., Smith, S., Parker, S., & Martin, R. (1997). Learning and retention of auditory warnings. In E. Mynatt & J. A. Ballas (Eds.), *Proceedings of the fourth International Conference on Auditory Display, ICAD '99* (pp. 129–134). Palo Alto, CA: ICAD.

Lippa, Y. (1996). A referential-coding explanation for compatibility effects of physically orthogonal stimulus and response dimensions. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49A,* 950–971.

Lu, C. -H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review, 2,* 174–207.

Marks, L. E. (1987). On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 384–394.

Martins, A. C. G., & Rangayyan, R. M. (1997). Experimental evaluation of auditory display and sonification of textured images. In E. Mynatt & J. A. Ballas (Eds.), *Proceedings of the fourth International Conference on Auditory Display, ICAD '99* (pp. 129–134). Palo Alto, CA: ICAD.

Martins, A. C. G., Rangayyan, R. M., Portela, L. A., Amaro, E., Jr., & Ruschioni, R. A. (1996). Auditory display and sonification of textured images. In S. Frysinger & G. Kramer (Eds.), *Proceedings of the third International Conference on Auditory Display, ICAD '96* (pp. 9–11). Palo Alto, CA: ICAD.

Melara, R. D., & Marks, L. E. (1990a). Dimensional interactions in language processing: Investigating directions and levels of crosstalk. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 539–554.

Melara, R. D., & Marks, L. E. (1990b). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics, 48,* 169–178.

Melara, R. D., & Marks, L. E. (1990c). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 398–414.

Melara, R. D., & O'Brien, T. P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General, 116,* 323–336.

Melara, R. D., & O'Brien, T. P. (1990). Effects of cuing on cross-modal congruity. *Journal of Memory and Language, 29,* 655–686.

Michaels, C. F. (1988). S–R compatibility between response position and destination of apparent motion. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 231–240.

Michaels, C. F. (1993). Destination compatibility, affordances, and coding rules: A reply to Proctor, Van Zandt, Lu, and Weeks. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1121–1127.

Moore, B. C. J. (1989). *An introduction to the psychology of hearing* (3rd ed.). London: Academic Press.

Mudd, S. A. (1963). Spatial stereotypes of four dimensions of pure tone. *Journal of Experimental Psychology, 66,* 347–352.

Neuhoff, J. G., & McBeath, M. K. (1996). The Doppler illusion: The influence of dynamic intensity change on perceived pitch. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 970–985.

Patterson, R. D. (1982). *Guidelines for auditory warning systems on civil aircraft.* CAA Paper No. 82017, Civil Aviation Authority, London.

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.

Proctor, R. W., & Reeve, T. G. (Eds.). (1990). *Stimulus–response compatibility: An integrated perspective.* Amsterdam: North Holland.

Proctor, R. W., Van Zandt, T., Lu, C. -H., & Weeks, D. J. (1993). Stimulus–response compatibility for moving stimuli: Perception of affordances or directional coding? *Journal of Experimental Psychology: Human Perception and Performance, 19,* 81–91.

Robinson, D., & Dadson, R. (1957). Threshold of hearing and equal-loudness relations for pure tones, and the loudness function. *Journal of the Acoustical Society of America, 29,* 1284–1288.

Roswarski, T. E., & Proctor, R. W. (1998). *Auditory stimulus-response compatibility: Is there a role for a motor code?* Manuscript in preparation.

Schiller, P. (1935). Interrelation of different senses in perception. *British Journal of Psychology, 25,* 465–469.

Schneider, W. (1988). Micro Experimental Laboratory: An integrated system for IBM-PC compatibles. *Behavior Research Methods, Instrumentation, & Computers, 20,* 206–217.

Schneider, W. (1995). *MEL Professional user's guide.* Pittsburgh, PA: Psychology Software Tools.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus–response compatibility: An integrated perspective* (pp. 31–86). Amsterdam: North Holland.

Simon, J. R., Mewaldt, S. P., Acosta, E., Jr., & Hu, J. (1976). Processing auditory information: Interaction of two population stereotypes. *Journal of Applied Psychology, 61,* 354–358.

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

Valenzuela, M. L., Sansalone, M. J., Krumhansl, C. L., & Streett, W. B. (1997). Use of sound for the interpretation of impact-echo signals. In E. Mynatt & J. A. Ballas (Eds.), *Proceedings of the fourth International Conference on Auditory Display, ICAD '99* (pp. 47–56). Palo Alto, CA: ICAD.

Walker, B. N., & Kramer, G. (1996). Mappings and metaphors in auditory displays: An experimental assessment. In S. Frysinger & G. Kramer (Eds.), *Proceedings of the third International Conference on Auditory Display, ICAD '96* (pp. 71–74). Palo Alto, CA: ICAD.

Walker, R. (1987). The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perception & Psychophysics, 42,* 491–502.

Wang, H., & Proctor, R. W. (1996). Stimulus–response compatibility as a function of stimulus code and response modality. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1201–1217.

White, B. W. (1969). Interference in identifying attributes and attribute names. *Perception & Psychophysics, 6,* 166–168.

Wier, C. C., Jesteadt, W., & Green, D. M. (1977). Frequency discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America, 61,* 178–184.

Received December 5, 1997
Revision received March 22, 1999
Accepted March 25, 1999 ■