Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: Tests of adult and infant listeners[a)]

E. Glenn Schellenberg[b)]
Department of Psychology, University of Windsor, Windsor, Ontario N9B 3P4, Canada

Laurel J. Trainor
Department of Psychology, McMaster University, Hamilton, Ontario L8S 4K1, Canada

Two experiments examined the influence of sensory consonance on the perceptual similarity of simultaneous pairs of complex tones (harmonic intervals). In experiment 1, adults heard a sequence of five consonant intervals (each a perfect fifth, or 7 semitones) and judged whether a subsequently presented test interval was a member of the sequence. Discrimination performance was better when the test interval was dissonant (tritone, 6 semitones) rather than consonant (perfect fourth, 5 semitones), despite the fact that the change in interval width was twice as great for the consonant than for the dissonant comparison. In experiment 2, 7-month-old infants were tested with an operant headturn procedure in a similar design and exhibited an identical pattern of responding. Hence, for both age groups, consonance was more important than interval width in determining the perceived similarity of harmonic intervals. © *1996 Acoustical Society of America.*

PACS numbers: 43.66.Hg, 43.66.Fe, 43.75.Cd [JWH]

INTRODUCTION

When adult listeners evaluate simultaneous pairs of complex tones (harmonic intervals), they typically judge intervals with small-integer frequency ratios to be *consonant*, or pleasant sounding; by contrast, larger-interval ratios are judged to be *dissonant*, or unpleasant (e.g., for a review, see Schellenberg and Trehub, 1994b). Nonetheless, judgments of similarities among intervals tend to be based on similarities in frequency distance (i.e., interval width) as much as they are on frequency ratios, or consonance. For example, in order to map intervallic similarities using multidimensional scaling, Levelt *et al.* (1966) asked their listeners to make similarity judgments of 15 different complex-tone intervals. The resulting dimensions indicated that similarity judgments were based on interval width as well as consonance. Indeed, our reanalysis of these judgments (Levelt *et al.*, 1966, Table 3) revealed that the proportion of variance in judgments explained by similarity in consonance (22.4%) was no different from that explained by similarity in width (26.3%).[1] Other researchers have reported that width can be more important than consonance in determining similarities among intervals. For example, when listeners identify harmonic intervals by name (e.g., major second, perfect fifth, etc.), errors tend to involve intervals similar in width rather than consonance (Plomp *et al.*, 1973).

The focus of the present investigation was on *sensory consonance* [also referred to as *tonal* (Plomp and Levelt, 1965) or *psychoacoustic* (Bregman, 1990) consonance] as opposed to *musical* consonance. Sensory consonance is a function of physical properties of the stimulus and is therefore independent of exposure to music or to cultural differences in musical styles. Accordingly, the responses of musically naive young infants are particularly informative in determining the extent to which natural factors affect perceptions of consonance. By contrast, musical consonance is considered to be a learned phenomenon resulting from exposure to music; combinations of tones that are musically consonant in one culture could be musically dissonant in another culture.

Sensory consonance is a function of *critical bandwidth* (Kameoka and Kuriyagawa, 1969; Plomp and Levelt, 1965), reflecting the fact that the human auditory system is unable to fully resolve tones that are proximate in pitch. For pure-tone intervals, sensory consonance is considered to be strictly a matter of tone proximity. Nonidentical pure tones that are proximate in frequency fall within the same critical band (frequency range) and can only be partially resolved by the basilar membrane; the resulting excitation patterns overlap and thus interact. Listeners perceive amplitude fluctuations—which Helmholtz (1885/1954) referred to as "roughness"—in such instances. When a pure-tone interval is sufficiently wide so that its component tones can be fully resolved, it is consonant because no roughness is perceived. Critical bandwidth is less than 3 semitones for frequencies higher than 500 Hz, with maximum roughness occurring at approximately one-quarter of a critical band (i.e., when tones are slightly less than 1 semitone apart); for lower frequencies, critical bandwidth is more or less a constant difference in frequency (about 80 Hz) rather than a constant ratio (Rasch and Plomp, 1982). Critical bandwidths correspond to equal distances along the cochlear partition (Greenwood, 1991) and do not differ substantially in size between infants and adults at any frequency (Schneider *et al.*, 1990; Olsho, 1985). Because perceptions of roughness are a physical phenomenon and independent of exposure to music, sensitivity

---

[a)]Presented at the 10th Biennial International Conference on Infant Studies.
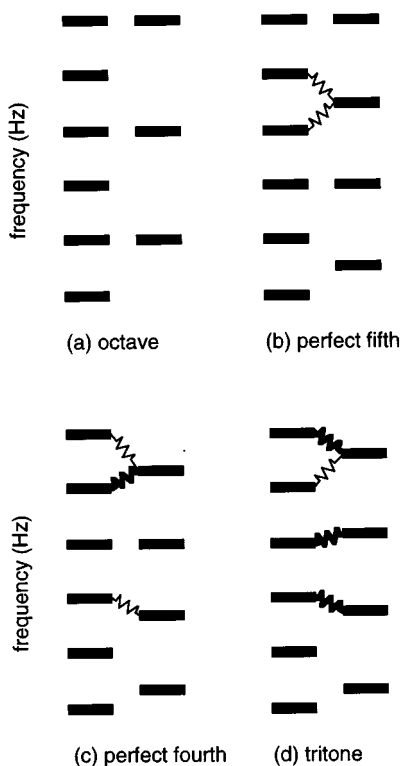[b)]Corresponding author.

FIG. 1. Schematic drawing of four pairs of complex tones: (a) an octave, (b) a perfect fifth, (c) a perfect fourth, and (d) a tritone. Ratios of fundamental frequencies (higher tone: lower tone) are 2:1, 3:2, 4:3, and 45:32 for (a)–(d), respectively. Black bars represent harmonics. For each pair, the first six harmonics of the lower tone are shown. Adjacent harmonics that fall within a critical band, causing roughness (sensory dissonance), are connected with a jagged line. Darker lines indicate instances of relatively extreme dissonance (i.e., ≤1 semitone between adjacent harmonics).

to differences in sensory consonance could be similar for all listeners regardless of age or cultural background.

For intervals comprised of two *complex* tones, any pair of adjacent harmonics can fall within a critical band (see Fig. 1). Compared to intervals with larger-integer ratios, complex-tone intervals with small-integer frequency ratios have more adjacent harmonics that are identical and relatively few that are nonidentical (Helmholtz, 1885/1954), so dissonance is less likely as the integers in a frequency ratio become smaller. For example, when two complex tones form an octave interval (2:1 ratio), all of the harmonics of the higher tone are also harmonics of the lower tone [see Fig. 1(a)], so octaves are free of dissonance. For complex tones a perfect fifth apart (3:2 ratio), evenly numbered harmonics of the higher tone are also harmonics of the lower tone. Because the third harmonic of the higher tone and the fourth harmonic of the lower tone fall within a critical band (they are 2 semitones apart), as do the third and fifth harmonics of the higher and lower tones, respectively (also 2 semitones apart), some roughness is perceived [see Fig. 1(b)]. The situation is similar for perfect fourths (4:3 ratio) [see Fig. 1(c)]. Thus perfect fifths and perfect fourths are more dissonant than octaves, although they are still considered *perfect consonances* in Western music theory. By contrast, complex-tone tritone intervals are considered dissonant, having a large-integer frequency ratio (45:32), *no* identical harmonics

that are audible, and very rough (1-semitone) intervals between several pairs of adjacent harmonics [see Fig. 1(d)].

In the present study, we examined whether listeners perceive two consonant intervals to be more similar than a consonant and a dissonant interval. We sought to determine whether listeners' sensitivity to consonance would influence their ability to discriminate one interval from another interval, the assumption being that perceptually similar intervals should be poorly discriminated but that dissimilar intervals should be more discriminable. Discrimination of a dissonant interval (tritone, 45:32 ratio) from a consonant interval (perfect fifth, 3:2) was expected to be better than discrimination of two consonant intervals (perfect fifth, 3:2, and perfect fourth, 4:3), even though the latter comparison involved a larger difference in interval width. Hence, the design pitted influences of interval width and consonance directly against each other.

To increase the likelihood that our listeners' responses would be independent of musical consonance, we attempted to limit the influence of learned musical relations by presenting randomly ordered stimuli in chromatic (nontonal) contexts to listeners selected without regard to musical training. If effects of learning were to persist despite such attempts, they should be considerably greater among adults (experiment 1)—with extensive informal exposure to music—than among 7-month-old infants (experiment 2). Although infants of this age have been exposed to music, knowledge of Western scale structure appears to be relatively undeveloped in infants both younger and older. For example, 6-month-old infants are equally likely to detect alterations to a tone sequence when it is composed with a Javanese scale (which contains intervals that are not present in Western music) or a Western major scale, whereas adults perform better with the Western sequence (Lynch *et al.*, 1990). Similarly, in contrast to 5-year-old children and adults, the ability of 8-month-old infants to detect alterations to a Western melody is independent of whether or not such alterations violate the melody's scale structure (Trainor and Trehub, 1992, 1994).

I. EXPERIMENT 1

In the present experiment, adult listeners' discrimination of a consonant and a dissonant interval was compared to their discrimination of two consonant intervals. Listeners judged whether a comparison interval was the same as (or different from) a standard interval. The method differed from conventional same–different (AB) tasks in the following way: On each trial, listeners heard *five* repetitions of the standard interval (i.e., a sequence of standards) before the comparison (test) interval was presented. The standard was a consonant interval (perfect fifth, or 7 semitones, frequency ratio of 3:2) presented in transposition (i.e., each of the five presentations differed in pitch). Listeners' task was to judge whether the test interval was a member of the sequence. The test interval was a perfect-fifth interval from the sequence, a dissonant interval (tritone, 6 semitones, 45:32 ratio) that was 1 semitone narrower than a perfect fifth, or a consonant interval (perfect fourth, 5 semitones, 4:3 ratio) that was 2 semitones narrower than a perfect fifth. If judgments in this context are influenced more by sensory consonance than by

differences in width, listeners should be superior at detecting the smaller change. Alternatively, if interval width is a stronger determinant of perceived similarities, a change of 2 semitones should be more reliably detected than a change of 1 semitone.

A. Method

1. Subjects

The listeners were 28 members of the University of Windsor community who received course credit or token remuneration for participating in the experiment, which took approximately 30 min. Listeners were recruited without regard to musical background; six had more than five years of music lessons ($M = 8$ years), the other 22 had five years or less ($M = 2.3$ years).

2. Apparatus

Stimulus presentation and response recording were controlled by a Power Macintosh 7100/66AV computer and a customized software program. The stimuli were presented to listeners with lightweight SONY CD550 headphones in a sound-attenuating booth manufactured by Excel Industries. Stimuli were musical instrument digital interface (MIDI) files created using the CUBASE 1.8.3 (Steinberg, Inc.) music sequencing software program. Stimuli were generated on-line by a Roland JV-90 expandable synthesizer connected to the computer with a MIDI interface (Mark of the Unicorn MIDI Express).

3. Stimuli

On each trial, listeners heard a sequence of five contiguous equal-tempered perfect-fifth (7-semitone) intervals, each of which was 250 ms, followed by a 750-ms period of silence and a 250-ms test interval (see Fig. 2). Although equal-tempered consonant intervals are not tuned to *exact* small-integer frequency ratios, deviations from such ratios are very small (i.e., 2 cents, or 2% of 1 semitone) for equal-tempered perfect fifths and perfect fourths. The intervals were presented with a digitally sampled piano timbre (Roland JV-90 factory preset: Acoustic Piano 11).

Each of the five perfect-fifth intervals in the sequence was a semitone apart [e.g., 261.6 and 392.0 Hz (C_4–G_4), 277.7 and 415.3 Hz ($C_4^\sharp$–$G_4^\sharp$), 293.7 and 440.0 Hz (D_4–A_4), 311.1 and 466.2 Hz ($D_4^\sharp$–$A_4^\sharp$), and 329.6 and 493.9 Hz (E_4–B_4)], such that the ten component tones of the sequence did not belong to any musical key. On *no-change* trials, the test interval was identical to the second highest interval in the preceding sequence [e.g., 311.1 and 466.2 Hz ($D_4^\sharp$–$A_4^\sharp$) for the sequence listed above; see Fig. 1]. Test intervals on change trials were also comprised of component tones from the sequence. On *1-semitone* trials, the top tone of the test interval was displaced downward by 1 semitone relative to the test interval of no-change trials, forming an interval of 6 semitones [tritone, e.g., 311.1 and 440 Hz ($D_4^\sharp$–A_4)]. On *2-semitone* trials, the top tone was displaced downward by 2 semitones, forming an interval of 5 semitones [perfect fourth, e.g., 311.1 and 392.0 Hz ($D_4^\sharp$–$G_4^\sharp$)].
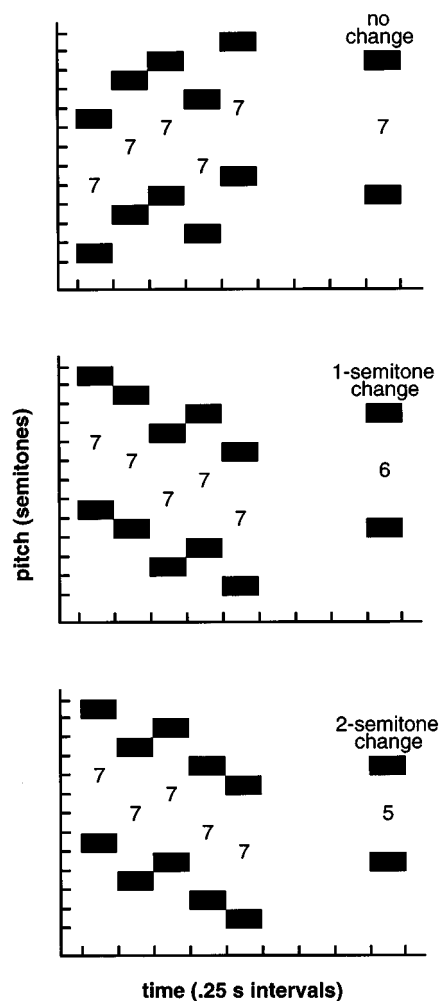

FIG. 2. Schematic drawing of the three types of trials used to test adult listeners in experiment 1. Numbers indicate interval width in semitones.

On each trial, sequence intervals were presented in random order, constrained such that the lowest interval was presented first and the highest interval last, or the highest interval first and the lowest interval last. This constraint was implemented to eliminate primacy or recency effects and ensured that component tones of test intervals were never components of the first or the last interval of the preceding sequence. The entire sequence was presented at one of three different pitch levels; the bottom tone of the lowest interval was 261.6 (C_4), 277.2 ($C_4^\sharp$), or 293.7 (D_4) Hz. Pitch level was selected randomly on each trial such that there was an equal number of trials at each level. Each of 12 possible sequence orders was used for each of the three types of trials (no-change, 1-semitone, 2-semitone) at each of the three pitch levels, for a total of 108 trials. The 108 trials were presented in a different random order for each listener.

To familiarize listeners with the procedure, two demonstration and eight practice trials were presented before the actual experiment began. The change to be detected was more obvious for these trials than for actual test trials. Specifically, the test interval of change trials was displaced upward or downward in pitch by an octave from the test interval of no-change trials.

E. G. Schellenberg and L. J. Trainor: Sensory consonance   3323

4. Procedure

Details of the procedure were provided to listeners both verbally and on the computer screen. Listeners viewed the computer screen through a window in the sound-attenuating booth and used a mouse to initiate trials and to record whether or not the test interval was a member of the preceding sequence (''yes'' or ''no''). They initially heard two demonstration trials (a change trial followed by a no-change trial) and eight practice trials (four change, four no-change). Following the practice trials, listeners were informed that discrepancies between the test interval and sequence intervals during the actual experiment would be more subtle than they were during practice trials. They then completed the experimental trials. After each practice and experimental trial, feedback (''correct'' or ''incorrect'') was provided on the computer screen.

B. Results and discussion

Overall performance was 67%, 75%, and 69% correct on no-change, 1-semitone change, and 2-semitone change trials, respectively. For each listener, two discrimination (d') scores (Elliott, 1964) were calculated. A 1-semitone score was derived using proportions of ''hits'' on 1-semitone trials (correctly responding ''no'') and ''false alarms'' (incorrectly responding on no-change trials). A 2-semitone score was derived similarly, substituting the hit rate with proportions of hits on 2-semitone trials. To avoid the possibility of infinite d' scores, proportions of hits and false alarms were transformed for each listener by adding 0.5 to the numerator (the number of ''no'' responses) and 1 to the denominator (the number of trials), following Thorpe *et al.* (1988). Infinite d' scores are thought to reflect sampling error due to the relatively small number of trials rather than ''perfect'' discrimination (Thorpe *et al.*, 1988). Although this transformation changes d' scores slightly, it does not alter their rank order. After transforming hit and false-alarm rates, the maximum d' score was 4.64.

Listeners' d' scores are illustrated in the scatter plot (Fig. 3). Preliminary analyses revealed that listeners' discrimination was well above chance levels ($d'=0$) for both comparisons [1-semitone change: $t(27)=5.84$, $p<0.0001$, $M=1.34$, s.d.$=1.21$; 2-semitone change: $t(27)=4.96$, $p<0.0001$, $M=1.11$, s.d.$=1.18$]. A paired t test indicated that listeners were significantly better at detecting the 1-semitone change than they were at detecting the 2-semitone change, $t(27)=2.48$, $p=0.02$ (difference score: $M=0.23$, s.d.$=0.50$). A nonparametric test examining differences in median levels of performance confirmed that the advantage for the 1-semitone change was consistent across listeners [Wilcoxon signed-ranks test (normal approximation) $z=2.27$, $p=0.02$]; 18 of 28 listeners performed better on the 1-semitone change, seven performed better on the 2-semitone change, and three performed equally well on both comparisons.

On one hand, these results replicate those reported earlier by Levelt *et al.* (1966), who found that subjective judgments of similarities among complex-tone intervals varied as a function of sensory consonance. On the other hand, judg-
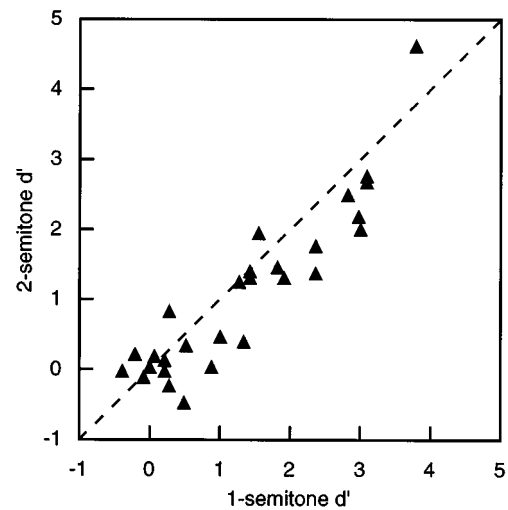

FIG. 3. Discrimination (d') scores for adult listeners in experiment 1. Points below the diagonal indicate better discrimination of 1-semitone (dissonant) changes than of 2-semitone (consonant) changes.

ments of Levelt *et al.*'s listeners were affected equally by differences in interval width. By contrast, listeners' performance in the present experiment was more influenced by similarities in consonance than by similarities in width, even though the difference in width was twice as great for the comparison between consonant intervals (7 and 5 semitones) as for the consonant–dissonant comparison (7 and 6 semitones).

The next group of analyses examined whether response patterns were influenced by musical consonance. We tested whether listeners' years of formal music lessons could predict their d' scores, the rationale being that if listeners' judgments were influenced by learned musical relations, those with more training in music would have been particularly likely to exhibit perceptual grouping based on intervallic consonance rather than width. The results revealed that neither 1-semitone nor 2-semitone scores were associated with musical training (r's<0.2, p's>0.3). To test whether musical training could predict the observed advantage for 1-semitone changes over 2-semitone changes, an ''advantage'' (i.e., difference) score was calculated for each listener by subtracting 2-semitone d' scores from 1-semitone d' scores; advantage scores were also uncorrelated with musical training, $r=0.23$, $p>0.2$.

Thus listeners' *formal* training in music had no effect on response patterns. Nonetheless, all of our adult listeners had years of *informal* exposure to Western music. Virtually all intervals with small-integer ratios (e.g., 2:1, 3:2, 4:3, 5:3, 5:4, 6:5) and little sensory dissonance are structurally important in Western music. Hence, implicit knowledge of familiar intervals garnered from informal exposure to music could be used to explain the present results. Relatively uncommon intervals may be perceptually distinct from common intervals simply because they are unfamiliar. By studying listeners with considerably less exposure to music, however, influences of consonance can be examined in a context where effects of familiarity would be minimized.
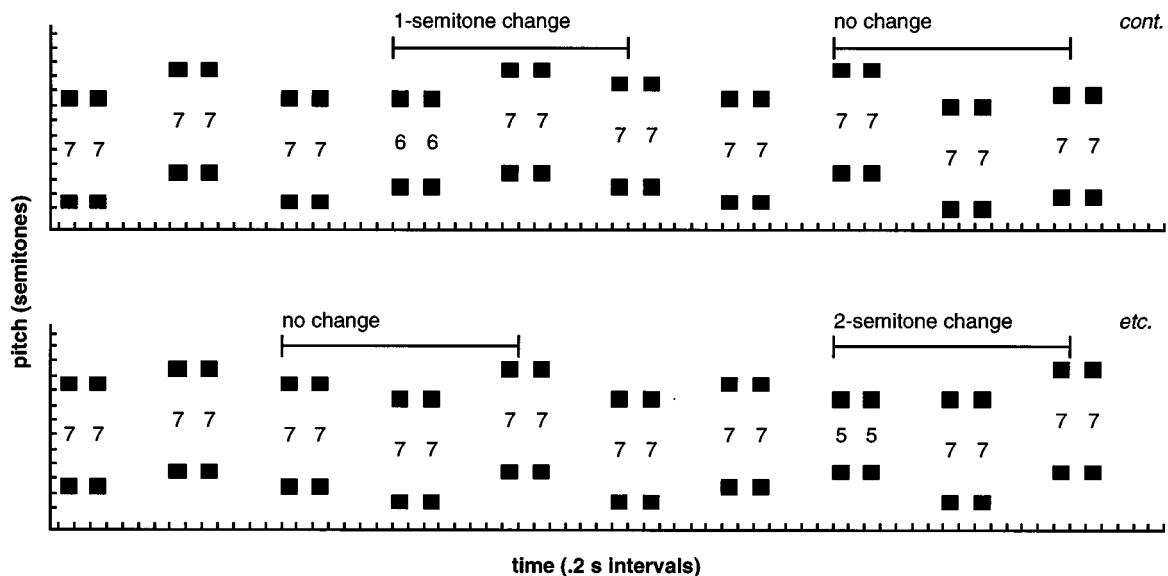
FIG. 4. Schematic drawing of the method used to test infant listeners in experiment 2, illustrating the three types of trials and the 3-s response window for each trial. Numbers indicate interval width in semitones.

II. EXPERIMENT 2

In the present experiment, 7-month-old infants were tested in a design similar to that of experiment 1. Because direct same/different judgments are not possible with infants, a variant of the operant headturn procedure (see Trehub and Trainor, 1993) was used. The young age of the listeners made it unlikely that exposure to music would be the source of their response patterns. Each infant heard a perfect-fifth interval (7 semitones) presented repeatedly at different pitch levels and was trained to turn toward a loudspeaker when the interval changed. Headturn responses were monitored during three types of trials: no-change trials (another presentation of a perfect fifth), 1-semitone change trials (presentation of a tritone, or 6 semitones), and 2-semitone change trials (presentation of a perfect fourth, or 5 semitones).

A. Method

1. Subjects

The listeners were 15 infants between 6.5 and 7.5 months of age ($M = 7$ months, 3 days) who were recruited from families living near McMaster University. All infants were born within 2 weeks of term; they weighed at least 2500 g at birth with no known abnormalities and were healthy at the time of testing. No infants were eliminated for crying or fussing.

2. Apparatus

The experiment was controlled by a Macintosh IIci computer. Stimulus tones were generated with Synthesize (a sound-generation software program) and an Audiomedia (DigiDesign) 16-bit sound card. A customized software program controlled stimulus presentation. An assistant used a button-box connected to a Strawberry Tree I/O card (via a custom-built interface) to initiate trials and to record responses. Stimuli were presented with a Denon amplifier

(PMA-480) and an audiological GSE loudspeaker. Testing was conducted in a sound-attenuating chamber (Industrial Acoustics Co.). Infants sat on a parent's lap facing the assistant. The loudspeaker was located 45° to the infant's left on top of a box with a smoked Plexiglas front that contained four compartments, each with a set of lights and a mechanical toy. During reinforcement for correct responding, the lights were illuminated in one of the compartments, which enabled the infant to see an activated toy. The assistant sat behind a small table that concealed the button box and had an assortment of hand puppets that were used to attract the infant's attention between trials.

3. Stimuli

As in experiment 1, the stimuli were harmonic intervals comprised of two complex tones. The tones consisted of the first ten harmonics added with random phases. Harmonics decreased successively in amplitude with a falloff of 6 dB per octave. Tones were 200 ms including 10-ms linear onsets and offsets. They were presented with an amplitude of 60 dB (A) at the approximate location of the infant's head.

The repeating background (standard) pattern consisted of two repetitions of a perfect-fifth interval (exact 3:2 frequency ratio) separated by 200 ms of silence. The standard pattern was presented at three different pitch levels, such that the fundamental frequencies of its component tones were 261.6 and 392.4 Hz (C_4–G_4), 277.2 and 415.8 Hz ($C_4^\#$–$G_4^\#$), or 293.7 and 440.5 Hz (D_4–A_4) (see Fig. 4). Successive patterns were always presented at different pitch levels (chosen randomly) and were separated by 800 ms of silence. No-change trials consisted of another repetition of the standard pattern (i.e., a pair of perfect fifths) and were therefore indistinguishable from the repeating background. During the test phase, component tones of change trials were also components of the repeating background. Specifically, on

1-semitone change trials, the standard pattern was replaced by a pair of tritone intervals (6 semitones, 45:32 frequency ratio, 277.2 and 389.8 Hz, $C_4^{\#}$–G_4), whereas on 2-semitone change trials, the standard pattern was replaced by a pair of perfect-fourth intervals (5 semitones, 4:3 frequency ratio, 293.7 and 391.6 Hz, D_4–G_4). During the training phase, change trials consisted of a pair of minor-second intervals (1 semitone), with fundamental frequencies of 293.7 and 313.2 Hz (16:15 frequency ratio, D_4–$D_4^{\#}$). This change of 6 semitones in interval width can be considered neutral with respect to the experimental hypothesis because it was both dissonant and had a large difference in width compared to the perfect-fifth standard.

4. Procedure

The caregiver and assistant listened to masking music through headphones. Infants sat on their caregiver's lap in the booth and faced the assistant. The background (standard) pattern was presented repeatedly from a loudspeaker situated 45° to the infant's left. The assistant manipulated hand puppets to attract the infant's attention. When the infant was facing directly forward, the assistant called for a trial by pressing a button on the button box. Although the number of background patterns between trials could vary according to the infant's looking behavior, the minimum was three. The assistant pressed another button on the button box whenever the infant turned toward the loudspeaker. The computer recorded only those headturns that occurred within 3 s after the onset of a trial (see Fig. 4). Hence, the response window began with the onset of a trial pattern (one pair of intervals) and continued over the next one and a half patterns. Correct headturns (i.e., during the response window for change trials) resulted in visual reinforcement, which consisted of illumination and activation of a mechanical toy for 2 s. Incorrectly turning toward the loudspeaker at any other time had no consequence.

The test phase consisted of 24 trials, with equal numbers of no-change, 1-semitone, and 2-semitone trials. Trials were presented in random order, constrained such that no more than two no-change trials could occur consecutively, which precluded the possibility of relatively long periods with no chance of reinforcement and subsequent loss of interest on the part of the infant. The test phase was preceded by a training phase, during which infants were trained to turn toward the loudspeaker when the repeating background (standard) pattern changed. The training phase was identical to the test phase with the following exceptions: (1) all trials were change trials, (2) there was only one type of change trial, and (3) change trials were more obvious than those in the test phase (see Sec. II A 3). After an infant made four correct consecutive headturns on change trials, the training phase was terminated and the test phase began.

B. Results and discussion

Infants' performance accuracy was 83%, 50%, and 38% on no-change, 1-semitone change, and 2-semitone change trials, respectively. The relatively low false-alarm rate (17%) and hit rates (50% and 38%) reveal an overall "conservative
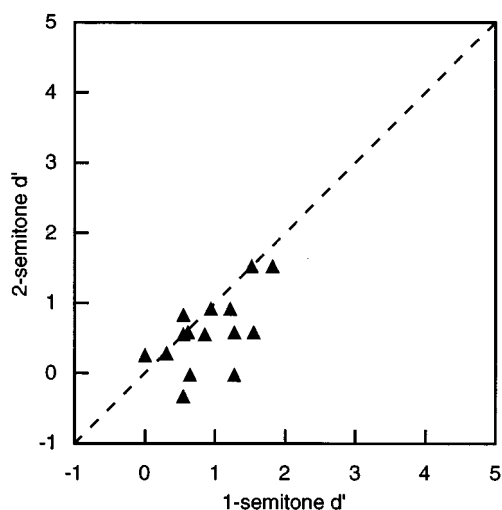

FIG. 5. Discrimination (d') scores for infant listeners in experiment 2. Points below the diagonal indicate better discrimination of 1-semitone (dissonant) changes than of 2-semitone (consonant) changes.

bias'' on the part of the listeners, which likely stemmed from their limited attention span. As in experiment 1, two d' scores were calculated for each infant (see Fig. 5). Comparisons with chance levels of performance revealed that infants successfully detected both the 1-semitone change, $t(14) = 6.85$, $p < 0.0001$ ($M = 0.91$, s.d. $= 0.52$), and the 2-semitone change, $t(14) = 4.49$, $p = 0.0005$ ($M = 0.60$, s.d. $= 0.52$). Infants were better at detecting the smaller 1-semitone change than they were at detecting the larger 2-semitone change, $t(14) = 2.56$, $p = 0.02$ (difference score: $M = 0.31$, s.d. $= 0.47$). The observed effect was *not* a consequence of the performance of a few infants: a Wilcoxon signed-ranks test (normal approximation) confirmed that median levels of performance were higher for the 1-semitone change than they were for the 2-semitone change, $z = 2.19$, $p = 0.03$. Eight infants exhibited an advantage for the 1-semitone change; only two showed the opposite pattern (five showed no difference). Hence, infants' overall pattern of responding was identical to that of adults with various degrees of musical training (experiment 1).

III. GENERAL DISCUSSION

Adults and 7-month-old infants were tested on their discrimination of changes in the width of a consonant complex-tone harmonic interval (perfect fifth, 7 semitones, 3:2 frequency ratio). Both groups of listeners were better at detecting a relatively small change in width (1 semitone) that resulted in a dissonant interval (tritone, 6 semitones, 45:32 ratio) than they were at detecting a change of twice the magnitude (2 semitones) that resulted in another consonant interval (perfect fourth, 5 semitones, 4:3 ratio). Hence, for both groups of listeners similarities of complex-tone intervals were based more on sensory consonance than they were on interval width.

Our findings conflict with those from earlier studies, which reported that the effect of interval width on intervallic similarity was equal to that of consonance (Levelt *et al.*,

1966) or even stronger (Plomp *et al.*, 1973). This discrepancy is likely methodological in origin. In general, listeners in auditory experiments may find it easier to attend to a single dimension on which stimuli vary. In contexts where interval widths vary greatly (i.e., from 1 to 16 semitones in Levelt *et al.*, 1966, from 1 to 12 semitones in Plomp *et al.*, 1973), listeners would be particularly likely to focus on differences in width. When width differences are minimized, however, as in the present experiments (i.e., from 5 to 7 semitones), influences of other dimensions would be more likely to emerge, such as those attributable to sensory consonance. Our use of multiple examples of the standard interval presented in transposition may have further helped to obscure differences in interval width and to make intervallic quality, or consonance, more noticeable.

Inevitable differences between the methods of experiments 1 and 2 (infants cannot make same/different judgments; adults would show ceiling effects if tested in the infant procedure) precluded direct comparisons between age groups. Nonetheless, patterns of responding among adults and infants were identical in kind if not in quantity, and the performance of adult listeners did not depend on their musical training. These findings raise the possibility, then, that sensitivity to sensory consonance and dissonance is independent of learning or exposure to music. Evidence of such sensitivity in other species—where exposure to music is not an issue—would provide support for this suggestion. For example, because birds perceive amplitude modulations much like humans do (Langner, 1992), they may also be sensitive to auditory roughness cues. Hulse *et al.* (1995) tested this hypothesis by presenting chords comprised of three simultaneous complex tones to European Starlings (*Sturnus Vulgaris*). The starlings were initially trained to peck at one key when a consonant chord (i.e., a major triad, e.g., C_4–E_4–G_4) was presented and at another key when a dissonant chord (e.g., C_4–D_4–G_4) was presented. Relatively rapid learning of novel pairs of stimuli revealed that the birds could generalize the distinction to a new pair of consonant (e.g., C_4–F_4–$G_4^{\#}$) and dissonant chords (e.g., C_4–$D_4^{\#}$–B_4). Thus sensitivity to sensory consonance could be a fundamental characteristic of auditory perception across species.

The results of the present study can also be interpreted within an alternative framework suggested by studies of the perception of *pure-tone* intervals. For example, infants are better able to detect changes to simultaneous pure-tone intervals that have small-integer rather than larger-interval frequency ratios (Schellenberg and Trehub, 1996b; Trainor, in press). Because component tones of intervals in these studies did not fall within a critical band, the results cannot be explained in terms of sensory consonance. Sensory consonance is also irrelevant to *melodic* intervals (sequential tone pairs). Nonetheless, infants appear to be sensitive to similarities between successive pure tones standing in a 2:1 (octave) frequency ratio (Demany and Armand, 1984). Moreover, studies of infants', children's, and adults' perception of sequential pure tones consistently show superior detection of alterations to such sequences when they feature small-integer rather than larger-integer frequency ratios (Schellenberg and Trehub, 1994a, 1996a, b; Trainor and Trehub, 1993a, b).

Regardless of sensory consonance, then, compared to intervals with larger-integer ratios, those with small-integer ratios appear to be encoded by listeners with relative ease, resulting in a stable perceptual representation and an enhanced ability to detect subtle alterations to such intervals.

Hence, perceptual similarities between intervals with small-integer frequency ratios could, conceivably, stem from their privileged perceptual status. It seems more likely, however, that a causal association would be in the opposite direction. Indeed, the present results are explained more simply in terms of sensitivity to sensory consonance. Specifically, we found that adults and infants are sensitive to the roughness cues that arise when tones can only be partially resolved. Such sensitivity could promote rapid learning of the frequency relations of tone combinations that are consonant, or naturally pleasant sounding (no roughness). This learning could then be generalized to other contexts, such as those involving pure tones presented simultaneously or sequentially. Similar to this hypothesis is Terhardt's (1978, 1984) suggestion that infants learn about consonance and dissonance through exposure to the overtone structure of vowels. Future research using discrimination measures to test the perceived similarity of pure-tone intervals (with no sensory dissonance) could determine more completely the source of the results reported here.

[1]Levelt *et al.* (1966, Table 1) provide similarity scores for 105 pairs of intervals (each pairwise combination of 15 different intervals). We reanalyzed these scores as a function of differences in width between the intervals in each pair (measured in semitones), and as a function of differences in consonance [measured using the index derived by Schellenberg and Trehub (1994b) to quantify the simplicity of a frequency ratio]. The correlation between differences in width and similarity scores ($r=-0.512$) was no different from the correlation between differences in consonance and similarity scores ($r=-0.473$), $t(102)=0.364$, $p>0.5$.

Bregman, A. S. (**1990**). *Auditory Scene Analysis* (Bradford, Cambridge, MA).

Demany, L., and Armand, F. (**1984**). ''The perceptual reality of tone chroma in early infancy,'' J. Acoust. Soc. Am. **76**, 57–66.

Elliott, P. B. (**1964**). ''Tables of *d′*,'' in *Signal Detection and Recognition by Human Observers: Contemporary Readings*, edited by J. A. Swets (Wiley, New York), Appendix 1, pp. 651–684.

Greenwood, D. D. (**1991**). ''Critical bandwidth and consonance in relation to cochlear frequency-position coordinates,'' Hear. Res. **54**, 164–208.

Helmholtz, H. L. F. von (**1954**). *On the Sensations of Tone as a Physiological Basis for the Theory of Melody*, edited and translated by A. J. Ellis (Dover, New York), 2nd ed. (Original work published 1885.)

Hulse, S. H., Bernard, D. J., and Braaten, R. F. (**1995**). ''Auditory discrimination of chord-based spectral structures by European starlings (*Sturnus Vulgaris*),'' J. Exp. Psychol.: Gen. **124**, 409–423.

Kameoka, A., and Kuriyagawa, M. (**1969**). ''Consonance theory: Part I. Consonance of dyads,'' J. Acoust. Soc. Am. **45**, 1451–1459.

Langner, G. (**1992**). ''Periodicity coding in the auditory system,'' Hear. Res. **60**, 115–142.

Levelt, W. J. M., van de Geer, J. P., and Plomp, R. (**1966**). ''Triadic comparisons of musical intervals,'' Br. J. Math. Stat. Psychol. **19**, 163–179.

Lynch, M. P., Eilers, R. E., Oller, D. K., and Urbano, R. C. (**1990**). ''Innateness, experience, and music perception,'' Psychol. Sci. **1**, 272–276.

Olsho, L. W. (**1985**). ''Infant auditory perception: Tonal masking,'' Infant Behav. Dev. **8**, 371–384.

Plomp, R., and Levelt, W. J. M. (**1965**). ''Tonal consonance and critical bandwidth,'' J. Acoust. Soc. Am. **38**, 548–560.

Plomp, R., Wagenaar, W. A., and Mimpen, A. M. (**1973**). ''Musical interval recognition with simultaneous tones,'' Acustica **29**, 101–109.

Rasch, R. A., and Plomp, R. (**1982**). ''The perception of musical tones,'' in *The Psychology of Music*, edited by D. Deutsch (Academic, New York), pp. 1–24.

Schellenberg, E. G., and Trehub, S. E. (**1994a**). ''Frequency ratios and the discrimination of pure tone sequences,'' Percept. Psychophys. **56**, 472–478.

Schellenberg, E. G., and Trehub, S. E. (**1994b**). ''Frequency ratios and the perception of tone patterns,'' Psychonom. Bull. Rev. **1**, 191–201.

Schellenberg, E. G., and Trehub, S. E. (**1996a**). ''Children's discrimination of melodic intervals,'' Dev. Psychol. **32** (in press).

Schellenberg, E. G., and Trehub, S. E. (**1996b**). ''Natural musical intervals: Evidence from infant listeners,'' Psychol. Sci. **7**, 272–277.

Schneider, B. A., Morrongiello, B. A., and Trehub, S. E. (**1990**). ''Size of critical band in infants, children, and adults,'' J. Exp. Psychol.: Human Percept. Perform. **16**, 642–652.

Terhardt, E. (**1978**). ''Psychoacoustic evaluation of musical sounds,'' Percept. Psychophys. **23**, 483–492.

Terhardt, E. (**1984**). ''The concept of musical consonance: A link between music and psychoacoustics,'' Music Percept. **1**, 276–295.

Thorpe, L. A., Trehub, S. E., Morrongiello, B. A., and Bull, D. (**1988**). ''Perceptual grouping by infants and preschool children,'' Dev. Psychol. **24**, 484–491.

Trainor, L. J. (**in press**). ''The effect of frequency ratio on infants' and adults' discrimination of simultaneous intervals,'' J. Exp. Psychol.: Human Percept. Perform.

Trainor, L. J., and Trehub, S. E. (**1992**). ''A comparison of infants' and adults' sensitivity to Western musical structure,'' J. Exp. Psychol.: Human Percept. Perform. **18**, 394–402.

Trainor, L. J., and Trehub, S. E. (**1993a**). ''Musical context effects in infants and adults: Key distance,'' J. Exp. Psychol.: Human Percept. Perform. **19**, 615–626.

Trainor, L. J., and Trehub, S. E. (**1993b**). ''What mediates infants' and adults' superior processing of the major over the augmented triad?,'' Music Percept. **11**, 185–196.

Trainor, L. J., and Trehub, S. E. (**1994**). ''Key membership and implied harmony in Western tonal music: Developmental perspectives,'' Percept. Psychophys. **56**, 125–132.

Trehub, S. E., and Trainor, L. J. (**1993**). ''Listening strategies in infancy: The roots of music and language development,'' in *Thinking in Sound: The Cognitive Psychology of Human Audition*, edited by S. McAdams and E. Bigand (Oxford U. P., London), pp. 278–327.