Psychology and Aging
1996, Vol. 11, No. 2, 235–246

0882-7974/96/$3.00

Effects of Aging and Musical Experience on the Representation of Tonal Hierarchies

Andrea R. Halpern
Bucknell University

SeYeul Kwak, James C. Bartlett, and W. Jay Dowling
University of Texas at Dallas

Two experiments explored the representation of the tonal hierarchy in Western music among older (aged 60 to 80) and younger (aged 15 to 22) musicians and nonmusicians. A probe tone technique was used: 4 notes from the major triad were presented, followed by 1 note chosen from the 12 notes of the chromatic scale. Whereas musicians had a better sense of the tonal hierarchy than nonmusicians, older adults were no worse than younger adults in differentiating the notes according to musical principles. However, older adults were more prone than younger adults to classify the notes by frequency proximity (pitch height) when proximity was made more salient, as were nonmusicians compared with musicians. With notes having ambiguous pitch height, pitch height effects disappeared among older adults but not nonmusicians. Older adults seem to have internalized tonal structure, but they sometimes fail to inhibit less musically relevant information.

Musical activities are popular among people of all ages. As in other areas of expertise, training and skill in music vary widely. However, music is unusual in that a larger proportion of the educated population has received at least some training in music, compared with other domains such as chess or art. Outside of formal training, exposure to music via the radio or concerts is enjoyed by large numbers of older people on a frequent basis (Gilbert & Beal, 1982). Indeed, among the hundreds of younger and older adults tested in our laboratory in recent years, 100% have answered *yes* when asked whether they enjoy listening to music. Gibbons (1982) surveyed 150 older adults and found that 126 desired to improve their musical skills during retirement by taking instrumental or singing lessons. So both passive and active participation in music is important to many older people.

Such widespread participation in music provides an ideal opportunity to study contributions of aging and specific training to cognitive skills, for several reasons. First, for some abilities such as face perception, most people can be considered experts in perceiving familiar facial types, and thus differential experience effects cannot easily be considered (Bartlett, 1993). In contrast, it is possible to find younger and older adults both with and without musical training, which allows the separation of these two factors. Second, musical training does not seem largely confined to one sex or the other, as might be the case in chess or needlework. Finally, music listening is an activity relatively impervious to cohort effects. Whereas listening preferences may vary across the generations, classical, big band, and rock music are all based on essentially the same principles of Western music.

The fact that music is so widely enjoyed even by people not specifically trained in the domain suggests that some aspects of music are learned implicitly. It seems hard to believe that people would expose themselves to patterns that made no sense to them. Music is an art based on recurring patterns in acoustic frequencies and time, and detection of patterns is essential to understanding music. We selected one of those regularities, the tonal hierarchy, and asked how well listeners maintain and access that knowledge into their 60s and 70s.

We begin by explaining the musical concept under discussion. An important structural constraint in Western music is the hierarchy of tonal relations. This means that notes in a scale serve different functions in music; they are not all equally important in a given context. To illustrate, consider sitting down at a piano and making up a simple, pleasant melody, beginning with the note C. A scale in the key of C begins with that note and encompasses all the white keys on a piano up to the next C. Your made-up melody is unlikely to contain many black keys, which are the notes outside of the key of C. You may use all the other white piano keys, but you will probably use some more than others. You will likely begin and end on the C, and use G and E more frequently than A or B. In other words, E and G are closer musically to the base note C (called the *tonic*) than the physically closer note B, which is just one piano key below the tonic. And a C# [sharp], which is the black note physically closest to the C (one semitone away) is quite distant musically from the tonic.

Andrea R. Halpern, Department of Psychology, Bucknell University; SeYeul Kwak, James C. Bartlett, and W. Jay Dowling, Program in Cognitive Science, University of Texas at Dallas. SeYeul Kwak and James C. Bartlett contributed equally to this project; their order of listing is arbitrary.

This research was supported by Grant R01-AG09965-01 from the National Institute on Aging. Portions of the research were presented at the November 1994 meeting of the Psychonomic Society in St. Louis, Missouri. We gratefully acknowledge the assistance of Paul Minda and Glenn Gesek in recruiting and testing participants.

Correspondence concerning this article should be addressed to Andrea R. Halpern, Department of Psychology, Bucknell University, Lewisburg, Pennsylvania 17837. Electronic mail may be sent via Internet to ahalpern@bucknell.edu.

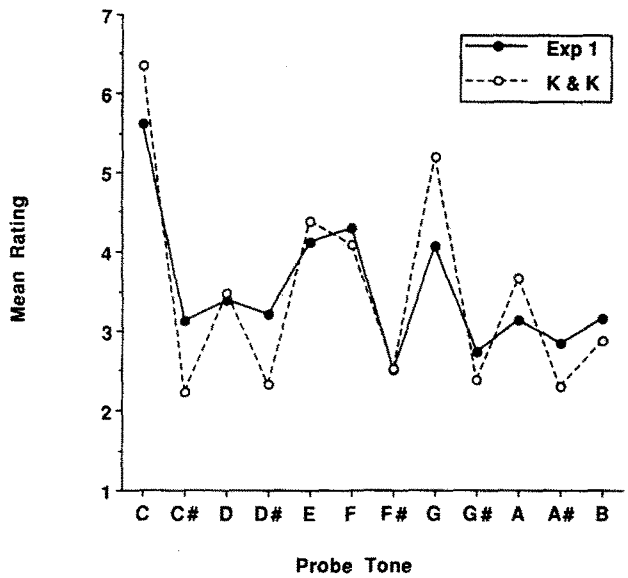

Figure 1. Mean goodness ratings (plus standard errors) for probe tones as endings for the musical context in the Experiment (Exp) 1, compared with data from Krumhansl and Kessler (K & K; 1982).

Musicians call these musical relations the *hierarchy of tonality.* Still using the key of C as our example, the C in the octave above is considered to be musically very close to the C in the lower octave. The notes E and G, which are the other members of the major triad, are considered next musically closest to the tonic. Other notes in the key, called *diatonic notes* (D, F, A, B) are considered the next closest, and notes outside the key (C#, D#, F#, G#, A#), called *nondiatonic notes,* are considered musically distant to the tonic. Note that this dimension of musical similarity is different from the dimension of frequency proximity, in which the closest notes to C would be, in ascending order, C#, D, D#, and so forth. The hierarchy of tonality is the basis of Western harmony and allows people to perceive points of movement and rest in melodies.

Whereas this notion of a hierarchy is derived through music theory, psychologists have shown that these relationships have psychological reality as well. A popular technique for demonstrating these relationships was devised by Krumhansl and Shepard (1979). In their *probe tone technique,* listeners hear a musical context followed by a probe tone, which can be any of the 12 notes of the octave (called the *chromatic scale*). Instructions are to rate the goodness of the probe tone as an ending to the context or sometimes to rate the goodness of the whole pattern. If listeners rate the goodness of the probe tone in the musically coherent order described above (highest rating to the tonic, then other triad tones, then other diatonic notes, then nondiatonic notes), researchers may infer that they have internalized the tonal hierarchy. Figure 1 (dashed lines) shows graphically the typical peak-and-valley pattern of responses consistent with the tonal hierarchy (listeners were young adults with some musical training; data are from Krumhansl & Kessler, 1982). The highest ratings were given to the C, E, and G, followed by the diatonic notes, followed by the nondiatonic notes, in conformance to the hierarchy described earlier.

Turning now to the link with cognitive aging, it hardly seems necessary to remind readers of this journal that some cognitive skills decline with age, particularly those requiring rapid access to newly learned information (see Hultsch & Dixon, 1990, for one among many recent reviews). However, the integrity of cumulative knowledge bases, for instance, the organization of semantic memory, seems unimpaired in normal aging (Light, 1988) even when access to that knowledge base may be slower among the older adults. Thus, vocabulary test scores are often higher among older than younger adults, whereas tip-of-the tongues states during explicit word retrieval are more common in older adults (Burke, MacKay, Worthley, & Wade, 1991), especially for less common words. Other tasks, such as lexical decisions, involve more implicit or automatic use of this lexical knowledge. In such cases, older adults are usually shown to be as sensitive as younger adults to aspects of the lexicon such as word frequency (Allen, Madden, Weber, & Groth, 1993; Tainturier, Tremblay, & Lecours, 1989), and semantic priming effects are often found to be at least as large among older adults as among younger adults (Duchek & Balota, 1993). These findings show that sensitivity to structural constraints, in this case lexical and semantic, remains robust into older age, even if explicit retrieval of this information may become less efficient.

Very little research has explored whether musical skills in general follow patterns of age-related increases, stability, or declines seen in other domains. In a recent study from our laboratory (Halpern, Bartlett, & Dowling, 1995), we asked listeners to judge whether two melodies were exact transpositions of one another. We found that age effects were more likely to appear when general perceptual skills were needed for success (the two tunes differed in overall pitch contour). Age effects were minimized when more specific musical knowledge was required (the two tunes had the same contour but different intervals). In the latter case, musical experience (defined as fewer than 2 years vs. at least 8 years of music lessons) was the important variable. Lynch and Steffens (1994) also found that older adults without formal music training were able to detect a mistuned note as well as young adults, as long as the sequence was drawn from a familiar Western tonal system. With tunes drawn from an unfamiliar musical system, young adults were superior to older adults.

These results imply that performance on tasks that are mainly dependent on well-practiced musical skills may not decline with age. Note that our interval abstraction task and the mistuning detection task in familiar tonal systems (Lynch & Steffens, 1994) both require access to presumably well-learned aspects of the listener's tonal system, and these are just the tasks showing few if any age differences.

These studies, although few in number and providing only indirect evidence for the question at hand, suggest that older people may maintain their base of musical knowledge into later years and make adequate use of it when the task demands that they do so. Deficits emerge to the extent that more general cognitive strategies or capacities are engaged. Note that in our study (Halpern et al., 1995), we found little support for the reduction of age differences among musicians. We concluded that younger age and greater experience convey independent types of advantages. This finding is consistent with recent studies by Morrow and colleagues with airplane pilots (Morrow, Leirer, & Altieri, 1992; Morrow, Leirer, Altieri, & Fitzsimmons, 1994). This group of researchers has found that experience can compensate for age only in tasks nearly identical to ones pilots would encounter in the cockpit, arguing against any generalized effect of experience in maintaining skills otherwise declining with age.

The current study was an attempt to more directly probe the knowledge of the tonal hierarchy possessed by musically trained and untrained younger and older adults. We assumed that this knowledge accrues across the lifetime via simple exposure to the musical idiom. Previous studies using the probe tone technique have shown increasing sophistication of knowledge about the tonal hierarchy from childhood to young adulthood and, to some extent, with training. For instance, in the original study by Krumhansl and Shepard (1979), some adults did not show the peak-and-valley pattern characteristic of musical responding. Instead, they rated pitches closest in pitch to the tonic as being better endings and pitches farther from the tonic as being worse endings, in essentially a monotonic function. This is referred to as a *pitch height effect* and reflects sensitivity to the psychophysical rather than musical relationships among the pitches. A similar result was found by Frankland and Cohen (1990) using latency measures. Interestingly, both sets of researchers observed that level of training was only a weak predictor of who would respond using a pitch height strategy.

Other researchers have found that, under some circumstances, even young children show evidence of being aware of a tonal hierarchy, but researchers have also found evidence of a developmental sequence. For instance, Speer and Meeks (1985) gave second and fifth graders ascending and descending scale contexts followed by a probe tone. Both age groups rated the tonic more highly than the other notes and preferred the triad to the other diatonic notes. A developmental trend was seen in that the second graders preferred the diatonic over the nondiatonic notes only in the descending context, whereas the fifth graders showed this preference in both contexts. Both groups also showed some evidence of a pitch height effect. Cuddy and Badertscher (1987) presented tones with ambiguous pitch height (called *Shepard tones*) to children 6 to 12 years of age and adults. Compared with ordinary tones, the pitch height ambiguity of Shepard tones makes pitch height less available as a response dimension. Cuddy and Badertscher found that with the possibility of responding to pitch height lessened, and in a strong tonal context of the major triad, all ages rated the tonic over the triad notes, followed by the other diatonic notes, and then the nondiatonic notes. They even preferred the G to the E within the major triad, which is characteristic of adult responding (see Figure 1). All age groups responded more musically to the major triad context and less musically to other contexts such as a scale or a diminished triad (the notes C, E♭ [flat], G♭, for instance). Cuddy and Badertscher concluded that children have abstracted the tonal hierarchy at least to some degree by the early elementary school ages.

The literature reviewed so far suggests that even people without formal training develop some notion of the tonal hierarchy at a young age. This intuitive knowledge becomes more differentiated with listening over time and perhaps with training. There is also evidence of a pitch height effect in some younger and less trained listeners. Because pitch height is a simple psychophysical dimension, it may be more available to listeners in whom the tonal hierarchy is less firmly established. It also is possible that listeners can apprehend tonal structure as well as pitch height but that responding on the basis of tonal structure information may require suppression of pitch height information.

Some researchers have suggested recently that age-related declines in a number of cognitive tasks may be characterized as a failure of inhibitory processes with age. Hasher and Zacks (1988) theorized that inefficient inhibitory processes will allow irrelevant information to enter into working memory. As an example, McDowd and Oseas-Kreger (1991) found older adults to be deficient in suppressing responses to previously relevant information in a negative priming paradigm, and Dempster (1992) reviewed evidence that a wide range of performance deficits in older people, children, and patients with frontal lobe damage may be traceable to failures of inhibition.

Thus in the current study, we were interested in whether knowledge about the musical hierarchy might be preserved or even increased in older age, or whether difficulties in inhibiting secondary information, in this case pitch height, might cause older people to adopt a pitch height response mode instead. If so, we expected to see age differences diminish if cues to pitch height were made less available by the use of Shepard tones. We were also interested to see under what circumstances experience might modify any age effects. We expected trained musicians to show at least a somewhat more differentiated tonal profile than nonmusicians, but on the basis of literature reviewed earlier, we did not expect to find Age × Experience interactions.

In Experiment 1 we used stimulus tones with ambiguous or well-defined pitch height to examine the effects of age and experience. In Experiment 2 we examined pitch height effects more thoroughly than we could in Experiment 1, as will be detailed below.

Experiment 1

Older and younger musicians and nonmusicians heard a musical context, followed by a probe tone that they rated for how good an ending it made for the context. Because we wanted to facilitate the eliciting of a tonal profile from all our participants, we used a strongly tonal context of a major triad. In this first experiment, we presented some trials in a naturalistic string orchestra timbre and other trials using tones with ambiguous pitch height (Shepard tones). We predicted that older people and untrained people would show rating patterns consistent with a tonal hierarchy (the tonic rated most highly, followed by the other triad notes, the other diatonic notes, and then the nondiatonic notes), although the patterns would perhaps not be as sharply defined as those of musicians. We predicted that these same groups might show evidence for a pitch height effect with the string timbre but that this tendency would be attenuated with the Shepard tones.

Method

Participants. In both studies, younger listeners (15 to 21 years old) were students from Bucknell University. Older listeners (60 to 80 years old) were residents from communities in central Pennsylvania or from the Dallas, Texas, area and were recruited through newspaper ads, senior citizens' centers, and musical organizations. In both experiments, we defined musical experience in terms of years of private instrument or voice lessons. Musicians were defined as having had at least 8 years of private lessons and nonmusicians as having had no more than 2 years of lessons. In the case of self-trained musicians, we considered anyone with professional performing experience to be in the highly experienced category. Younger listeners were uncompensated volunteers or received course credit for participation. Older listeners received a small stipend for participation.

In Experiment 1, 12 people were recruited for each age–experience category. Mean age of the younger listeners (11 women, 13 men) was 19.0; mean age of the older listeners (sex distribution unavailable) was 70.2. Younger people had an average of 13.8 years of formal education, and older musicians had an average of 15.9 years of formal education (mean years of formal education for older nonmusicians was unavailable).

Materials. All tones were computer generated before being recorded onto cassette tape. The string orchestra tones were produced by a Yamaha PSR-510 synthesizer. They had a pleasant, rich quality. Shepard tones reduce the salience of pitch height by including many different octaves of a particular note in each tone. The lower and higher octaves in the tone are reduced in intensity relative to the middle octaves. If a scale of Shepard tones is played, the pitch appears to rise until about the middle of the scale. At this point, the loudest octaves have become softer and the previously less salient lower octaves have become louder. This produces an auditory analog of the always-ascending-staircase visual illusion. The scale note of Shepard tones is unambiguous, but the octave the pitch is played in is unclear.

Our particular Shepard tones were produced on a Zenith 386 microcomputer played through a single channel digital–analog converter at a sample rate of 10 kHz with a 16-bit quantization. Tones were low-pass filtered (4.2 kHz cutoff frequency). All tones contained five sine waves from five different octaves. The amplitude envelope was the same one used by Krumhansl, Bharucha, and Kessler (1982), where the middle octaves have the same amplitude, and amplitude decreases in a sinusoidal fashion for the extreme octaves. See Krumhansl et al. (1982) for more details of the construction of Shepard tones. The timbre thus produced had a slight raspy quality.

On each trial, a major triad context was sounded first. This consisted of the first, third, fifth, and then a repeat of the first notes of the scale. Tones sounded for 500 ms, with 250 ms between the tones. After a pause of 750 ms, the probe tone was sounded, which could be any of the 12 possible notes of the chromatic scale. In this experiment, probe tones between the first and fifth notes of the scale were always in the same octave as the context; however, probe tones outside this range were sometimes in the higher and sometimes in the lower octave compared to the context. A pause of 4 s separated each trial.

Four different keys were used in the experiment: C, A, D#, and F#. The key and probe tone of each trial were determined randomly, subject to the constraint that each key and each probe tone was sampled equally often in the session. Presentation of string tones or Shepard tones was blocked, and the order was counterbalanced over listeners. Each block contained 48 trials (each of the 12 chromatic probe tones, in each of four keys). Each session began with the practice trials described below.

Procedure. An experimental session began with administration of a musical background questionnaire. This was followed by a brief audiometric screening for our older listeners who did not wear hearing aids, to ensure that they had age-normal hearing. Using a Lucas GSI portable audiometer (Audio Electronics Inc., Houston), pure tone thresholds were measured from 250 Hz to 8,000 Hz. All participants then received the second half of the Wechsler Adult Intelligence Scale (Wechsler, 1981). We routinely administer this task to obtain a measure of cognitive functioning in an area unrelated to the musical tasks under investigation. Because older people typically exceed younger people on this task, any age differences in favor of younger people on the main tasks would be hard to attribute to general cognitive impairment.

Instructions for the probe tone task were given next. Participants were told that they would hear a sequence of four tones, followed by one more tone. Their task was to rate "how well the last tone fits with the rest of the music." Listeners were given a 7-point response scale, where 1 meant *absolutely bad ending* and 7 meant *perfect ending.* Intermediate points were labeled appropriately. Participants were further instructed to answer quickly and to trust their first impressions. They were asked to leave no blank spaces on the answer sheet.

Two sample sequences were played next, one of which illustrated a probe musically close to the tonic (the tonic itself) and one of which illustrated a musically distant probe (a nondiatonic note). Next, 10 sample trials were presented, after which participants were encouraged to ask any questions about the task. The two blocks of 48 experimental trials then ensued, separated by a brief rest period. All tapes were played on a high-quality stereo cassette player and speakers, and listeners could adjust volume to their liking. The session, including debriefing, lasted about 1 hr.

Results

Vocabulary scores. Out of a maximum score of 40, mean vocabulary scores of younger nonmusicians and musicians, followed by older nonmusicians and musicians were 20.0, 26.1, 29.8, and 27.1 respectively. Scores of older (M = 24.0) and younger (M = 28.5) participants differed significantly, $F(1, 44) = 7.55$, $p < .01$, $MSE = 45.9$, but scores of nonmusicians (M = 24.9) were not different from those of musicians (M = 26.6). There was an interaction, $F(1, 44) = 5.00$, $p < .05$; younger musicians had larger vocabulary scores than younger nonmusicians, but older musicians and nonmusicians had equivalent scores.

Probe tone ratings. This experiment had four variables: the two between-group variables of age (younger, older) and experience (musician, nonmusician) and the two within-group variables of instrument (string, Shepard tones) and probe tone. For data analysis, ratings from each key were averaged for each of the 12 probe tones and normalized to the key of C (in other words, a probe tone of C represents average ratings of the tonic presented in each of the four keys). Figure 1 shows the ratings for each probe tone averaged over the other variables. For comparison, the figure displays the probe tone ratings gathered by Krumhansl and Kessler (1982).

This global view of the data shows a peak-and-valley pattern similar to other studies in the literature. The notes of the major triad (C, E, G) and the F are rated the most highly, followed by the other diatonic notes. Nondiatonic notes received the lowest rating. The tonal profile is not as sharply differentiated as the comparison pattern. In particular, our listeners on average did not rate the D and A as highly as the listeners in Krumhansl and Kessler's experiment. However, our listeners were more heterogeneous in age and musical background than Krumhansl and Kessler's, and we used slightly different stimuli and response scales. Despite the differences in some aspects of results, the patterns are essentially similar, and the two patterns are significantly correlated, $r(10) = .92$.

We performed all our analyses in several ways. First, we treated probe tone as a 12-level variable in an analysis of variance (ANOVA). Second, we divided the tones into four tone types: the tonic (C), the other triad notes (E and G), the other diatonic notes (D, F, A, B), and the nondiatonic notes (C#, D#, F#, G#, B#). We mainly discuss this latter analysis, as the analysis using four tone types reflects the tonal hierarchy in a simpler way than dealing with all 12 probe tones. Specifically, listeners responding using the tonal hierarchy were expected to rate the tonic most highly, followed by the other triad notes, the other diatonic notes, and then the nondiatonic notes. We point out the interesting differences in results between the 4-level and 12-level analysis.

In addition, we performed all analyses using *z* scores as well as raw scores. For the *z* score analysis, each participant's scores were normalized using that person's own mean and standard deviation. This reduced variability associated with individual differences in using the response scale. For ease of presentation, we mostly report the raw score analysis in the text and the figures, as very few results were different in the two analyses. However, we report both analyses in the few cases when they were discrepant.

Effects not involving tone were of only secondary interest in this study because means may plausibly be interpreted as reflecting liking for the sequences in general. For completeness, we report that older people liked the sequences ($M = 4.13$) more than younger people ($M = 3.92$), $F(1, 352) = 5.22$, $p < .05$, *MSE* for all 4-level analyses = .77. This was modified by a trend toward older people preferring string tones over Shepard tones, whereas younger people showed no preference. This was a significant effect in the 12-tone analysis, $F(1, 1056) = 8.38$, $p < .01$, *MSE* for all 12-level analyses = 1.15, and in the *z* score analysis, $F(1, 352) = 5.97$, $p < .05$, $MSE = .34$. A similar trend, significant in the 12-tone analysis, shows that musicians preferred string tones, whereas nonmusicians had no preference, $F(1, 1056) = 5.89$, $p < .01$. In neither analysis did age and experience interact with each other or with any other variable.

We turn now to the variables of primary interest: the effects involving tone type in the 4-level analysis. A main effect shows that, as predicted, the tonic was rated most highly ($M = 5.61$), followed by the other triad notes, the other diatonic notes, and the nondiatonic notes ($Ms = 4.10, 3.49, 2.89$, respectively), $F(3, 352) = 170.03$, $p < .01$. Two 2-way interactions with tone type were significant. As can be seen in the top panel of Figure 2, tone type interacted with experience. Musicians differentiated the tone types more than nonmusicians, $F(3, 352) = 18.65$, $p < .01$. However, even nonmusicians showed the expected ordering of means for listeners responding on the basis of tonal hierarchies.

The bottom panel of Figure 2 shows an interaction of age and tone type. Although both older and younger listeners were differentiating the tone types in the expected fashion, the older listeners showed the more differentiated profile, $F(3, 352) = 5.05$, $p < .01$. However, this interaction was not significant in the *z* score analysis.

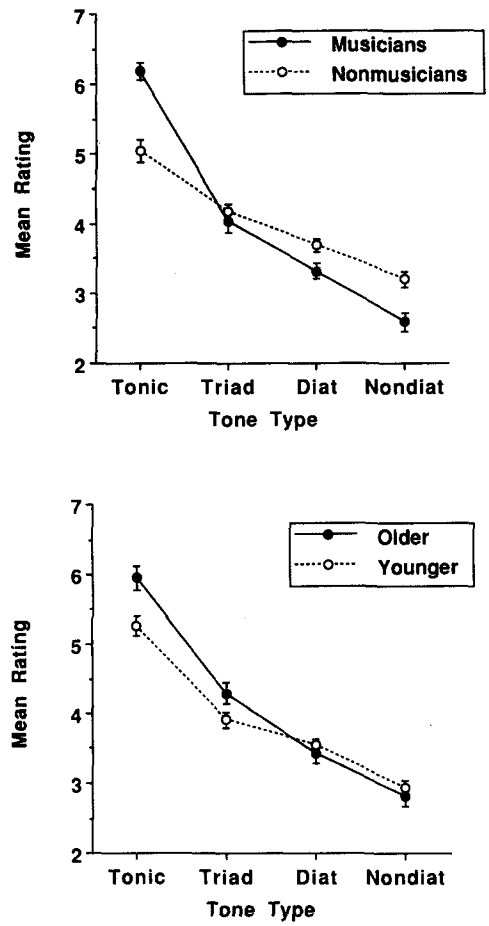

Figure 2. Mean goodness ratings (plus standard errors) in the interaction of experience and tone type in Experiment 1, showing greater differentiation of the tonal hierarchy by musicians versus nonmusicians (top) and older listeners versus younger listeners (bottom). Triad refers to triad notes other than the tonic. Diat = diatonic; nondiat = nondiatonic.

No other results were significant. In particular, there was no modification of the results of interest depending on type of instrument used. Outside of an overall preference for the string tones, especially among older people and musicians, the type of instrument did not affect the extent to which any of the listeners differentiated the tone types according to the tonal hierarchy. Even in the 12-level analysis, type of instrument appeared in no interactions with tone.

We also saw no evidence for pitch height effects. The rating profile for each type of listener looked essentially the same as that in Figure 2, with some variation in the extent to which the peaks were differentiated from the valleys. We did not see any functions that looked like a monotonic decrease of average rating with pitch distance from the tonic. However, as we discuss shortly, Experiment 1 may not have provided the best test of pitch height effects.

Discussion

To summarize, we presented a strongly tonal context of a major triad plus the tonic note, followed by a probe note. Young and old listeners, as well as musicians and nonmusicians, all rated the probe notes in a pattern consistent with using knowledge of the tonal hierarchy to decide on ratings. Tonic notes were rated the most highly, followed by other triad notes, followed by diatonic notes, followed by nondiatonic notes.

We were not surprised that musicians produced a more differentiated tonal profile than nonmusicians. In addition to more experience actually performing, musicians probably exceed nonmusicians in the amount of music they hear and the formal knowledge they have of tonal systems. We were, however, impressed with the extent to which the nonmusicians still demonstrated considerable knowledge of tonal relations. This knowledge was abstracted over the particular pitches being probed because everyone heard four different keys in the course of a session.

Rather than finding age-related impairments in this task, we found that the older listeners more strongly differentiated the tonic from the other triad notes and from the other diatonic notes than younger listeners in the raw score analysis. However, we must be cautious about this conclusion because of the failure to find this interaction in the z score analysis. A conservative interpretation is that a lifetime of listening can maintain knowledge of the tonal hierarchy, even if music has not been studied formally. Again, we note that both older and younger listeners showed the musical response pattern. We deliberately instantiated a musical context that conveyed a strong tonal sense. These circumstances apparently facilitate responding using a musical strategy, relative to some studies cited earlier where untrained adults did not always respond musically (e.g., Krumhansl & Shepard, 1979).

As predicted, we found no evidence for Age × Experience interactions in any analysis. Whereas the power of this interaction to detect a medium sized effect is low at .34 (following Cohen, 1969), we set $\alpha = .10$ and defined a medium effect as $d = .25$. This result is consistent with our theoretical arguments that age and experience generally have separate effects on cognitive tasks and with findings by ourselves and Morrow cited earlier (Halpern et al., 1995; Morrow et al., 1992, 1994). One interpretation worthy of further study is that long exposure to music or formal training is sufficient for displaying good tonal knowledge in these circumstances. In less optimal test conditions, such as hearing a less strongly tonal context, experience might very well confer larger advantages than age.

In this context it is interesting to note another way in which older age and musicianship seem to be similar. Both older listeners and musicians preferred the string timbre over the Shepard tones, whereas younger listeners and nonmusicians showed no preference. Whatever knowledge that accrues via explicit or implicit musical training may also encourage the development of an aesthetic sense that is more attracted to the harmonics found in normally experienced music than to the unusual construction of the Shepard tones.

We found no evidence for pitch height effects in this study, but as previously mentioned, Experiment 1 was not optimally suited to finding them. To keep the range of probe tones tested approximately the same across key, we varied the octave of probe tones with respect to the context in the different octaves. For instance, the probed tonic was always the identical note as the tonic in the triad, but the probed fifth was the same note as the fifth in the context in two of the keys but was an octave lower in the other two keys. In Experiment 2, we regularized the relationship of the probed notes to the tonic so that the probed notes were always higher than the last note of the context.

We were somewhat surprised to find that the type of instrument had no modifying influence on the extent to which tone types were differentiated (power to detect a medium sized difference = .90). Of course, the absence of a pitch height effect precluded finding an amelioration of that effect with Shepard tones. However, if younger listeners and nonmusicians were being influenced by pitch height even to a small extent, we might have expected some sharpening of the musical profile with the Shepard tones among those groups, as Cuddy and Badertscher (1987) found. One possibility for this lack of effect is that the string tones might have been perceived as being sufficiently musical to encourage attention to solely the musical closeness of the probe tone to the context. In Experiment 2 we compared Shepard tones with sine waves in an attempt to exaggerate the difference in the salience of pitch height.

Experiment 2

Experiment 2 was a replication of Experiment 1 with two modifications. First, all probe tones were drawn from the octave above the tonic in the musical context. This had the effect of regularizing pitch height relationships across the various keys. It also meant that the probed tonic was not the same note as the tonic in the context but was an octave higher. This allowed a stronger test of the primacy of the tonic in the ratings task than in Experiment 1. There, listeners may have rated the tonic highly because the identical note was sounded twice in the immediately preceding context. In Experiment 2, the tonic was the farthest note in frequency terms from the musical context. If listeners still rated the tonic as the best ending for the context, that would mean that they were responding to the musical closeness between the tonic sounded in the context and the tonic an octave higher rather than simply rating highly a note that was sounded twice previously.

The second modification was using sine waves instead of the string timbre. Sine waves by definition have no harmonics and thus convey very little information besides the note name and its octave. The comparison between sine waves and Shepard tones might thereby be a stronger manipulation of pitch height salience.

Our main prediction was that despite the musical responding demonstrated by older listeners in Experiment 1 with string timbres, we would find evidence of a pitch height effect when they were presented with sine waves. We expected the pitch height effect to become attenuated when pitch height was made an ambiguous cue by the use of Shepard tones. Once again, we predicted that, overall, musicians would show more evidence of using tonal relationships in their ratings than would nonmusicians but that age and experience would not interact in any of our results.

Method

Participants. Twelve new people meeting the criteria listed earlier were recruited for each age–experience category. Mean age of the youn-

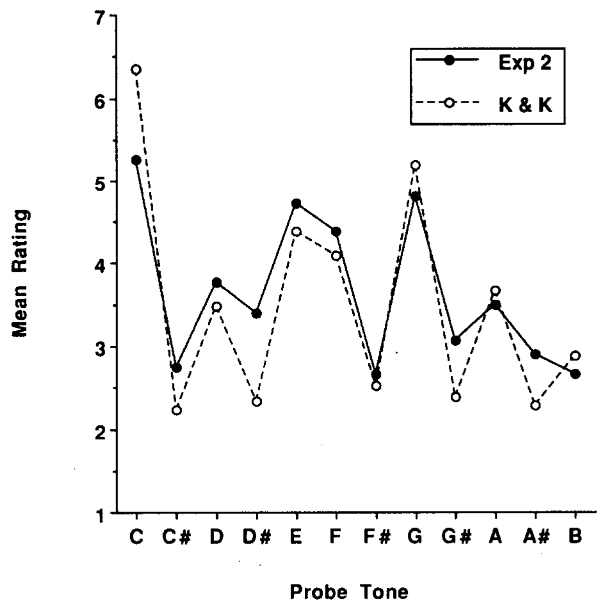

Figure 3. Mean goodness ratings for probe tones as endings for the musical context in Experiment (Exp) 2, compared with data from Krumhansl and Kessler (K&K; 1982).

ger listeners (16 women, 8 men) was 21.8 years; mean age of the older listeners (17 women, 7 men) was 67.7 years. Younger people had on average 15.9 years of formal education, and older people had on average 14.5 years.

Materials. Shepard tones were created in the same way as in Experiment 1. Sine waves were produced with a Yamaha TG-500 synthesizer. Onset and offset intensity envelopes were imposed on the sine wave voice so as to be as steep as possible without producing audible clicks on the loudspeakers.

Tapes were constructed in the same way as in Experiment 1, with the exception that all probe notes were higher in pitch than the final note of the context. Otherwise, timing and randomization characteristics were the same as in Experiment 1.

Procedure. Experiment 2 was conducted in the same way as Experiment 1. Shepard tones and sine waves were presented in blocks, with order of the blocks counterbalanced over listeners. All other procedures, including administration of the various subsidiary tasks, were the same as in Experiment 1.

Results

Vocabulary scores. Mean vocabulary scores of younger nonmusicians and musicians, followed by older nonmusicians and musicians, were 21.3, 21.8, 20.5, and 29.8, respectively. Scores of older and younger participants did not differ significantly, but scores of musicians (M = 25.8) exceeded those of nonmusicians, (M = 20.1), $F(1, 44) = 6.00$, $p < .05$, MSE = 46.7. There was an interaction, $F(1, 44) = 4.83$, $p < .05$, such that older musicians had larger vocabulary scores than older nonmusicians, but younger musicians and nonmusicians had equivalent scores.

Probe tone ratings. The first analysis of probe tone ratings was conducted as in Experiment 1. Figure 3 shows the overall rating pattern, again plotted with Krumhansl and Kessler's (1982) ratings for comparison. The correspondence between the two patterns appears to be even closer than in Experiment 1, and once again the patterns were significantly correlated, $r(10) = .93$.

As before, we performed analyses looking at all 12 probe tones and also divided the probe tones into the tonic, other triad notes, other diatonic notes, and nondiatonic notes. We report the 4-tone type analysis except where otherwise noted. We performed analyses on raw scores and z scores and again report the z score analysis only when it conflicts with the raw score analysis.

Older people (M = 4.33) liked the sequences more than younger people, (M = 3.95), $F(1, 352) = 13.75$, $p < .001$, MSE = 1.00 for all 4-level comparisons, although this preference was actually reversed in the z score analysis, $F(1, 352) = 4.90$, $p < .01$, MSE = .40. In the 12-tone analysis, older people liked sine waves better than Shepard tones, whereas younger people showed no preferences, $F(1, 1056) = 8.13$, $p < .001$, MSE = 1.34 in 12-tone analysis.[1] When experience groups were considered, musicians preferred the sine waves, whereas nonmusicians showed no preference, $F(1, 352) = 5.79$, $p < .05$. In no analysis did age and experience interact with each other or with any other factor.

Turning now to the variables of primary interest, the main effect of tone type was significant. Mean ratings for the four tone types, starting with the tonic, were 5.27, 4.77, 3.58, and 2.95, respectively; $F(1, 352) = 108.80$ $p < .001$. Similar to Experiment 1, the top panel of Figure 4 shows that musicians differentiated the tone types more than nonmusicians, $F(3, 352) = 32.23$, $p < .001$. The most notable difference between the two groups, and when compared with Experiment 1, was that nonmusicians here did not rate the tonic as the most preferred ending. The tonic and other triad notes were rated similarly, followed by the other diatonic and then nondiatonic notes (recall that here the tonic was psychophysically the most distant probe note from the context). An Experience × Instrument × Tone Type interaction with z scores only, $F(3, 352) = 3.95$, $p < .01$, indicates that this pattern was only evident for nonmusicians when they were listening to sine waves.

The bottom panel of Figure 4 shows that both older and younger people differentiated the tone types in accordance with the tonal hierarchy. We did not find an interaction of age with tone type in Experiment 2 with raw scores, but we did with z scores, $F(3, 352) = 7.39$, $p < .001$. The pattern of the interaction with z scores indicates that younger people showed more differentiation between other triad notes and diatonic notes than did older people.

The only remaining significant effect was the interaction of instrument and tone type illustrated in Figure 5, $F(3, 352) = 4.61$, $p < .01$. Participants differentiated between the tonic and

[1] Main effects not involving tone or tone type may differ in the 4-level and 12-level analyses because of the different weights placed on each level of this variable. For instance, in the 12-level analysis, ratings given to the tonic carry less weight than in the 4-level analysis. In the former case, the tonic is 1 of only 12 levels, whereas in the latter, it is 1 of 4 levels.

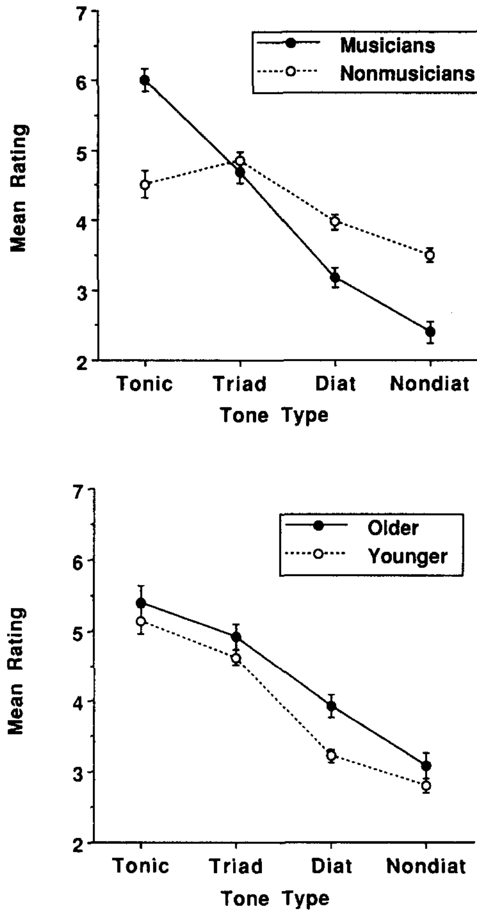

Figure 4. Mean goodness ratings (plus standard errors) in the interaction of experience and tone type in Experiment 2, showing greater differentiation of the tonal hierarchy by musicians versus nonmusicians (top) but similar differentiation by older versus younger listeners (bottom). Triad refers to triad notes other than the tonic. Diat = diatonic; nondiat = nondiatonic.

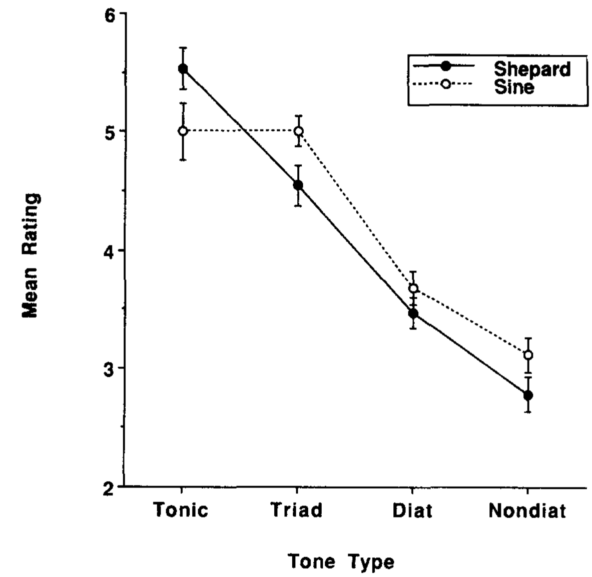

Figure 5. Mean goodness ratings (plus standard errors) in the interaction of instrument and tone type in Experiment 2, showing greater differentiation of the tonal hierarchy elicited by Shepard tones versus sine waves. Diat = diatonic; nondiat = nondiatonic.

other triad notes when listening to Shepard tones but rated them equivalently when listening to sine waves. Differentiation of other triad tones, diatonic tones, and nondiatonic tones did not seem to be affected by type of instrument.

Pitch height analysis. To analyze possible pitch height effects, we used data from 9 of the 12 probe tones. As shown in Table 1, we considered three levels of tone type: triad notes (including the tonic), diatonic notes, and nondiatonic notes. Within each level, a note was considered *close* if it was 2 to 4 semitones from the context tonic, *medium* if it was 7 to 9 semitones from the context tonic, and *far* if it was 10 to 12 semitones from the context. This enabled a factorial analysis of pitch height and tone type. For instance, an A is a diatonic note that is a medium distance from the tonic. Note that due to the stimulus construction, the C was an octave above the context, so that it is classified as a triad note that is far from the context. Close, medium, and far notes are only labeled relative to the context, not on an absolute pitch basis. Because four different keys were used, far notes were not necessarily the highest pitched notes in the overall pitch set, and close notes were not necessarily the lowest pitched notes in the overall pitch set.

We entered pitch height into an ANOVA, along with the other variables of age, experience, tone type (which is of course defined somewhat differently than in previous analyses), and instrument. We report only those effects involving pitch height, as other effects were essentially the same as reported previously. $MSE = 1.36$ for all F ratios.

First, the main effect of pitch height was significant. Mean ratings were 4.09, 3.79, and 3.60 for close, medium, and far pairs, respectively, $F(2, 792) = 12.86, p < .001$. Though statistically significant, all effects involving pitch height were small relative to effects involving tone type. The most and least preferred tone types differed by about 2.5 rating points in both experiments, whereas in Experiment 2 there was a difference of about 0.5 rating points between the close and far pairs.

Pitch height interacted with experience, as shown in Table 2,

Table 1
Probe Notes Used in the Pitch Height Analysis of Experiment 2

	Tone type		
Height	Triad	Diatonic	Nondiatonic
Close (2–4 semitones)	E	D	D#
Medium (7–9 semitones)	G	A	G#
Far (10–12 semitones)	C	B	A#

Note. Musical notes are labeled for the key of C. In the experiment, four different keys were used.

Table 2
Mean Probe Tone Ratings for Significant Interactions of Pitch Height With Experience Groups, Instruments, and Tone Types in Experiment 2

	Pitch height					
	Close		Medium		Far	
Variables	*M*	*SE*	*M*	*SE*	*M*	*SE*
Experience						
Musicians	3.75	0.12	3.48	0.14	3.52	0.18
Nonmusicians	4.42	0.09	4.10	0.11	3.67	0.11
Instrument						
Shepard tones	3.93	0.11	3.53	0.12	3.64	0.15
Sine waves	4.24	0.11	4.05	0.13	3.56	0.14
Tone type						
Triad	4.73	0.11	4.81	0.12	5.27	0.15
Diatonic	3.77	0.13	3.50	0.15	2.65	0.12
Nondiatonic	3.77	0.13	3.06	0.14	2.89	0.12

$F(2, 792) = 4.34, p < .01$. Musicians did not show any effect of pitch height, whereas nonmusicians rated the close notes higher than the medium notes, which they rated higher than the far notes (confirmed by tests of simple effects). Instrument interacted with pitch height, also shown in Table 2, $F(2, 792) = 5.08, p < .01$. Sine waves elicited higher ratings of close than medium than far notes, but Shepard tones did not elicit this pattern. Whereas close tones were rated most highly, the order of the next two tones was reversed (and all differences were rather small). The last entries of Table 2 show that tone type also interacted with pitch height, $F(4, 792) = 16.13, p < .001$. Closer inspection of Tables 1 and 2, however, shows that the lack or apparent reversal of a pitch height effect for triad notes was very likely due to the ordering of those notes in the tonal hierarchy. Specifically, the close note, E, was the least preferred of the triad notes and the far note, C, was of course the tonic, which elicits very high preference ratings. So for triad notes considered separately, a reversal in the pitch height effect was confounded with the musical preference ordering.

Because of this relationship between tone type and pitch height, the interpretation of two 3-way interactions involving these two variables is somewhat problematic. Tone type and pitch height interacted with experience, $F(4, 792) = 6.35, p < .001$. Inspection of the means suggest that the interaction was due to the particularly high rating given to the tonic by musicians, in an exaggeration of the Tone Type × Pitch Height interaction described above. Tone type and pitch height also interacted with age, $F(4, 792) = 3.86, p < .01$. Although both older and younger listeners rated the tonic very highly, older listeners also rated the A (the medium diatonic note) higher than the younger listeners. We have no explanation for this anomalous finding.

The one remaining significant effect is the most important one for our current purposes. As predicted, age and instrument interacted with pitch height, $F(2, 792) = 4.78, p < .01$. As shown in Figure 6, younger people (the dashed lines) showed only weak evidence of a pitch height effect. Older people listening to sine waves (filled circles, solid line) showed a strong pitch height effect. But these same listeners, when exposed to Shepard tones (open circles, solid line), showed no pitch height effect. (All these were confirmed by tests of simple effects.)

Discussion

Several results from Experiment 1 were replicated. Generally speaking, all listeners once again rated notes differently depending on their placement in the tonal hierarchy. As before, musicians showed a more differentiated tonal profile than nonmusicians. Particularly notable were the equivalent ratings given to the tonic and the other triad notes by the nonmusicians. This result could be due to the pitch height effect shown by nonmusicians as revealed in our second analysis. In this experiment, the probe tone tonic was musically close to the context tonic but far from it in terms of frequency. Nonmusicians may have been sensitive to both these dimensions, which would have resulted in a lower goodness rating of the tonic than the musicians, who were attending only to musical closeness.

Also similar to Experiment 1, we found no Age × Experience interactions in any analyses. We also found in both experiments that musicians and older people resembled one another in preferring the sine waves over the Shepard tones, whereas the other two groups showed no preference.

Unlike in Experiment 1, we found that the type of instrument affected some results. The Instrument × Tone Type interaction confirmed previous findings in the literature that, relative to timbres where pitch height is more salient, Shepard tones facilitate responding on the basis of the tonal hierarchy. Here, the

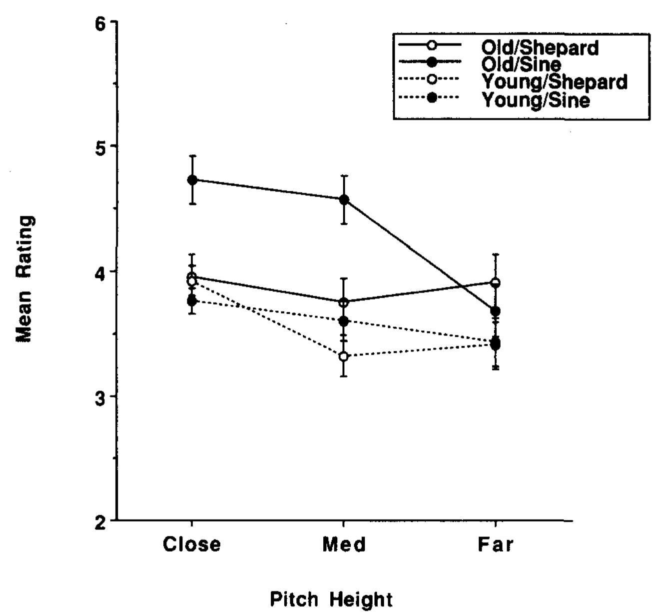

Figure 6. Mean goodness ratings (plus standard errors) in the interaction of age, instrument, and pitch height in Experiment 2, showing the strongest pitch height effect among older people listening to sine waves, which was eliminated when they listened to Shepard tones. The younger people showed weak pitch height effects. Med = medium.

tonic, sounded as a sine wave, was clearly musically close but psychophysically far from the context, perhaps attenuating the high preference rating usually given to the tonic. With Shepard tones, the pitch height of the tonic was ambiguous, apparently encouraging listeners to attend to the musical proximity dimension and rate the tonic highly.

By analyzing the relevant subset of our data, we found evidence of pitch height effects among some listeners, although they were small in magnitude relative to the effects of tone type. Consistent with the Instrument × Tone Type interaction just discussed (Table 2), pitch height effects were found more strongly with sine waves. The use of Shepard tones allowed listeners to successfully focus more on the tonal relationships between context and probe and not on the secondary dimension of pitch height. We also found that the less trained listeners were prone to pitch height effects, whereas the musicians were immune to them.

Our main prediction was upheld in the Age × Instrument × Pitch Height interaction (Figure 6), for both raw scores and *z* scores. The pitch height effect seems to be most prominent among the older listeners hearing sine waves. Sine waves make pitch height very salient, and to some degree the older listeners were "captured" by this dimension. But when the competing information was made difficult to use by presentation of the Shepard tones, the older listeners used their knowledge of the tonal hierarchy in formulating their ratings.

In Experiment 2, we did not find that the older listeners differentiated tone types more completely than the younger listeners (power to detect a medium sized effect = .64), unlike the raw score analysis of Experiment 1 (but similar to the *z* score analysis). In fact, the *z* score analysis in Experiment 2 suggests that younger people were differentiating some of the tone types more than the older people. To help clarify this discrepant pattern of results, we analyzed the Age × Tone Type interactions with *z* scores in both experiments using Shepard tones only. The reasoning for this is that we found evidence for pitch height effects with sine waves for older people in Experiment 2, where pitch height was well defined. By analyzing only Shepard tones, we could more clearly see if youth confers an advantage in differentiating the tone types once pitch height effects are removed. Indeed, we found no significant interaction of Age × Tone type in either experiment for Shepard tones (*MSE* for these *F* ratios = .30 and .37, for Experiments 1 and 2, respectively). The only Age × Tone Type interaction was found in Experiment 2 with sine waves, $F(3, 176) = 6.14$, $p < .001$, consistent with our argument that the older people were being affected by pitch height only in that condition.

General Discussion

Do listeners maintain their knowledge of the tonal hierarchy into their 60s and 70s? Our overall answer must be yes. Both experiments showed that older people differentiated their probe tone ratings according to the tonal hierarchy at least as well as younger people, once we removed the effect of pitch height. Because of the success of older musicians and nonmusicians in responding according to the tonal hierarchy, we conclude that mere exposure to music does allow people to abstract important regularities in their musical idiom, even if they cannot articulate that knowledge formally.

The exact nature of the knowledge conveyed by probe tone profiles is the subject of some disagreement among music psychologists. As summarized by Krumhansl (1990), probe tone profiles of the type presented in Figure 1 may reflect a truly abstract knowledge of tonal relationships (this is the way we presented the paradigm in the introduction). However, it is also possible that higher ratings are given to notes not because they are musically close to the key center but because they are more frequently sounded in music encountered by the listener over the short or long term. In fact, notes rated highly in tonal profiles do tend to appear more frequently in music. These two positions are difficult to disentangle, in that more frequent notes in a piece of music might very well have stronger tonal relationships to the tonic. The relevant point for the current project is that an important structural aspect of music is learned implicitly by young adulthood and is maintained, or perhaps even elaborated, by continued musical listening into older adulthood.

As mentioned earlier, we constructed our task so as to maximize the probability that all listeners would display a differentiated tonal profile by using a major triad context. Other musical contexts elicit less differentiated tonal profiles. For instance, Cuddy and Badertscher (1987) found some evidence for pitch height effects when children heard a scale context and found a nearly flat tonal profile when they heard a diminished triad context (the notes C, E♭, G♭). West and Fryer (1990) gave listeners a context composed of the seven scale notes, but randomly ordered. Both music students and nonmusic students rated the triad probe notes higher than the other notes but did not differentiate any further. An interesting question is, how would older listeners fare in these contexts? On the one hand, our current results predict that older listeners would be as able or even more able to abstract tonal relationships in these contexts than younger listeners, if attending to context–probe relationships relies mainly on musical listening skills. However, to the extent that extraction of tonal structure demands more from working memory or attentional capacity, age differences in favor of younger adults might be exaggerated. On this reasoning, we would expect older adult listeners to show a flatter musical profile than younger adult listeners when tonal relationships are harder to fit with the context.

At the same time that we found evidence for stability of tonal knowledge among our older listeners, we also found that they were more apt than younger listeners to use pitch proximity as a response mode in their ratings. Pitch height effects, although small in magnitude in our task, were evident most strongly among the older listeners when listening to sine waves. This is consistent with work cited earlier showing a relative deficiency in older adults in inhibiting irrelevant information. In a musical context, "correct" and "incorrect" answers and "relevant" and "irrelevant" information are less strictly defined than in some other domains. Nevertheless, if we assume that a musical task should elicit musically based responding, then the older listeners were letting information less relevant to this response affect them to some degree. The older listeners were able to use their musical knowledge perfectly well when the less relevant pitch height information was made less salient with the Shepard tones.

An interesting extension to this work might be to vary the instructions in the probe tone task to stress musical responding to a greater or lesser degree. If older people have some difficulty in suppressing the irrelevant dimension of pitch height, then instructions to use musical intuitions in answering, or to imagine the task as being relevant to composing, might help them to avoid using pitch height in their answers. If not, we could conclude that attention to a less relevant source of information is more hard-wired and less amenable to being influenced by shifts in strategy. It might also be illuminating to collect latencies to classify probe tones as good or bad endings. Janata and Reisberg (1988) and Frankland and Cohen (1990) found that the probe tones eliciting higher ratings in the standard paradigm also elicited faster classifications in a speeded version. If older listeners have an intact knowledge base of tonal information as well as facile retrieval from it, we can expect the same pattern of results from older and younger people. However, if older people have some trouble accessing the knowledge base, we can expect their latency pattern to be flatter than that of younger people, even if the rating profiles look the same.

We found no evidence for Age × Experience interactions in any tasks or analyses. Older musicians were not especially advantaged in differentiating the tonal hierarchy, and older nonmusicians were not especially prone to pitch height effects. Before concluding that age and experience fail to interact for substantive reasons, we must consider whether the lack of the interaction is merely due to ceiling effects or statistical reasons. First, were musicians and older people already differentiating the probe tones maximally so that no improvement was possible among older musicians? This seems unlikely given that the highest and lowest values were not at the endpoints of the rating scale (see Figures 2 and 4; note that the ordinate in both figures begins at 2 instead of 1).

Second, we noted earlier that although the power of the Age × Experience interaction is quite modest, converging evidence suggests that age and training are qualitatively independent variables. Evidence supportive of this position includes the fact that in both experiments, the nonmusicians showed a flatter tonal profile than the musicians, whereas this was not true of older people relative to younger people. Further, whereas the pitch height effect in older people was ameliorated by the Shepard tones in Experiment 2 (as shown by a significant Age × Instrument × Pitch Height interaction, Figure 6), we failed to find evidence that the pitch height effect in nonmusicians was significantly diminished by Shepard tones (no interaction of experience, instrument, and pitch height; power to detect a medium sized effect = .59). The notion that age and training are qualitatively different was also supported in our previous study (Halpern et al., 1995), in which we found differing effects of age and experience in the transposition recognition task, as described earlier.

In conclusion, the current work extends previous research in the verbal domain showing that knowledge bases may remain stable or even become more elaborated with age. In retrieval of this information, older people and nonmusicians seemed more prone to use psychophysical information than younger people and musicians. The use of this information by older people may be a manifestation of a general failure of inhibiting irrelevant information, as they (but not nonmusicians) benefited from a context in which the nonmusical information was made less salient. However, the small size of the pitch height effect relative to the generally well-differentiated tonal profiles makes it easy to understand why so many older people make listening to music an important part of their lives.

References

Allen, P. A., Madden, D. J., Weber, T. A., & Groth, K. E. (1993). Influence of age and processing stage on visual word recognition. *Psychology and Aging, 8,* 274–282.

Bartlett, J. C. (1993). Limits on loss in face recognition. In J. Cerella, J. Rybash, W. Hoyer, & M. L. Commons (Eds.), *Adult information processing: Limits on loss* (pp. 352–379). San Diego, CA: Academic Press.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30,* 542–579.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* San Diego, CA: Academic Press.

Cuddy, L. L., & Badertscher, B. (1987). Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. *Perception and Psychophysics, 41,* 609–620.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 12,* 45–75.

Duchek, J. M., & Balota, D. A. (1993). Sparing activation process in older adults. In J. Cerella, J. Rybash, W. Hoyer, & M. L. Commons (Eds.), *Adult information processing: Limits on loss* (pp. 383–406). San Diego, CA: Academic Press.

Frankland, B., & Cohen, A. J. (1990). Expectancy profiles generated by major scales: Group differences in ratings and reaction time. *Psychomusicology, 9,* 173–192.

Gibbons, A. C. (1982). Musical skill level self-evaluation in noninstitutionalized elderly. *Activities, Adaptation, and Aging, 3,* 61–67.

Gilbert, J. P., & Beal, M. R. (1982). Preferences of elderly individuals for selected musical education experiences. *Journal of Research in Music Education, 30,* 247–253.

Halpern, A. R., Bartlett, J. C., & Dowling, W. J. (1995). Aging and experience in the recognition of musical transpositions. *Psychology and Aging, 10,* 325–342.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–226). San Diego, CA: Academic Press.

Hultsch, D. F., & Dixon, R. A. (1990). Learning and memory in aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 259–274). San Diego, CA: Academic Press.

Janata, P., & Reisberg, D. (1988). Response-time measures as a means of exploring tonal hierarchies. *Music Perception, 6,* 161–172.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* New York: Oxford University Press.

Krumhansl, C. L., Bharucha, J. J., & Kessler, E. (1982). Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception & Performance, 8,* 24–36.

Krumhansl, C. L., & Kessler, E. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89,* 334–368.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception & Performance, 5,* 579–594.

Light, L. L. (1988). Language and aging: Competence versus performance. In J. E. Birren & V. L. Bengston (Eds.), *Emergent theories of aging* (pp. 177–213). New York: Springer.

Lynch, M. P., & Steffens, M. L. (1994). Effects of aging on processing of novel musical structure. *Journal of Gerontology: Psychological Science, 49,* P165–P172.

McDowd, J. M., & Oseas-Kreger, D. M. (1991). Aging, inhibitory processes, and negative priming. *Journal of Gerontology: Psychological Science, 46,* P340–P345.

Morrow, D. G., Leirer, V. O., & Altieri, P. A. (1992). Aging, expertise, and narrative processing. *Psychology and Aging, 7,* 376–388.

Morrow, D. G., Leirer, V. O., Altieri, P. A., & Fitzsimmons, C. (1994). When expertise reduces age differences in performance. *Psychology and Aging, 9,* 134–148.

Speer, J. R., & Meeks, P. U. (1985). School children's perception of pitch in music. *Psychomusicology, 5,* 49–56.

Tainturier, M., Tremblay, M., & Lecours, A. R. (1989). Aging and the word frequency effect: A lexical decision investigation. *Neuropsychologia, 27,* 1197–1203.

Wechsler, D. (1981). *WAIS-R Manual: Wechsler Adult Intelligence Scale—Revised.* New York: Psychological Corporation.

West, R. J., & Fryer, R. (1990). Ratings of suitability of probe tones as tones after random orderings of notes of the diatonic scale. *Music Perception, 7,* 253–258.

Received March 30, 1995
Revision received August 24, 1995
Accepted August 24, 1995 ■