Psychology and Aging
1995, Vol. 10, No. 3, 325-342

Aging and Experience in the Recognition of Musical Transpositions

Andrea R. Halpern
Bucknell University

James C. Bartlett and W. Jay Dowling
University of Texas—Dallas

The authors examined the effects of age, musical experience, and characteristics of musical stimuli on a melodic short-term memory task in which participants had to recognize whether a tune was an exact transposition of another tune recently presented. Participants were musicians and nonmusicians between ages 18 and 30 or 60 and 80. In 4 experiments, the authors found that age and experience affected different aspects of the task, with experience becoming more influential when interference was provided during the task. Age and experience interacted only weakly, and neither age nor experience influenced the superiority of tonal over atonal materials. Recognition memory for the sequences did not reflect the same pattern of results as the transposition task. The implications of these results for theories of aging, experience, and music cognition are discussed.

Research from the past 20 years has provided much evidence that the accuracy and efficiency of cognitive processing are related to age as well as experience. However, we still do not know how age-associated differences are related to experience effects. Researchers have shown that practice improves task performance in older adults and young adults (e.g., Salthouse & Somberg, 1982), and it has also been found that older persons perform as well as young adults in some tasks that are relevant to well-practiced skills (Charness, 1979, 1981; Salthouse, 1984). Nonetheless, three basic questions about aging and experience effects have yet to be adequately addressed.

The first question is whether experience reduces the size of age differences such that a highly trained older person can perform just as well as a highly trained young person in a domain-relevant task. A problem accompanying research on this question is that age and experience are often confounded: Older experts may have had much more practice than young adult experts in the same skill domain (Salthouse, 1990). This problem notwithstanding, if age-related differences in a cognitive task are reduced among those with much experience, one can conclude that experience (a) maintains or preserves abilities that would otherwise decline (the "use it or lose it" hypothesis), (b) allows older persons to strategically accommodate to their cognitive deficits (e.g., to maintain their speed in typing by looking further ahead on the page), (c) allows the task to be performed using relatively age-invariant processes (e.g., automatic processes), or (d) compensates somehow for declining information processing efficiency (e.g., by allowing older adults to base their responses on activation of prestored knowledge as opposed to computation). Regardless of which interpretation is valid, reduced age differences among people highly experienced in a domain would be a finding of theoretical and practical importance. Surprisingly, this finding has seldom been obtained (see Salthouse, 1990, for a general discussion; see Morrow, Leirer, & Altieri, 1992, for a recent example from the field of aviation).

The second question resembles the first except that it focuses on task materials instead of prior practice: Are age-related differences in a cognitive task increased when the stimuli are poorly structured, providing a poor "fit" to participants' knowledge and skills? Some findings suggesting that the answer is *yes* come from experiments on processing of language. For example, age-related deficits in memory for prose and for perception of time-compressed speech are exacerbated when the materials are linguistically poorly structured (Smith, Rebok, Smith, Hall, & Alvin, 1983; Wingfield, Poon, Lombardi, & Lowe, 1985). However, all normal adults are expert language users. We wish to expand this second question by asking whether the Age × Material interactions become more or less extreme among those persons more skilled in a domain than among relative novices.

The third question is whether any age-related deficits we do find in a domain-specific task reflect functional losses in skill or, rather, efficiencies or slow-downs in processing that impair task performance but are extrinsic to experience effects. To use Chomsky's (1965) well-known terms, do age-related differences in a field of experience reflect losses in *competence* within the domain, or simply losses in general *performance* capacities involved in many tasks? At an empirical level, the question becomes whether deficits in performance that are shown by older persons as compared with young adults are equivalent to the deficits shown by untrained people as compared with those more highly trained, or, alternatively, whether age-related differences and experience effects are qualitatively distinct and thus influenced by different aspects of the participants, tasks, or materials.

Andrea R. Halpern, Department of Psychology, Bucknell University; James C. Bartlett and W. Jay Dowling, Cognitive Science Program, University of Texas—Dallas.

This research was supported by National Institute on Aging Grant 1-R01-AGO9965-01. Portions of the research were presented at the November 1992 meeting of the Psychonomics Society in St. Louis, Missouri.

We gratefully acknowledge the assistance of the following students in preparing stimuli, recruiting and testing participants, and analyzing data: Melinda Andrews, Sharon Feldt, Julene Johnson, Se Yeul Kwak, Scott Lipscomb, Michael Rauso, and Jean Searcy. We thank Jackson Hill for help in preparing Figure 1.

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.

All three of these questions are important theoretically, as differing conceptions of cognitive aging and cognitive skill imply different answers. Consider, for example, the familiar intuition that age-related differences in cognitive performance reflect deterioration in neural "hardware," whereas experience effects reflect improvements in processing strategies or "software" (see Charness, 1985, p. 243). This hardware–software notion implies one answer to our third question: Age-related effects and experience effects should show qualitative differences. However, this notion makes no clear predictions about the other two questions.

Now consider the alternative view (Charness, 1985, p. 253) that aging is related to functional deficits in working memory, whereas domain-specific experience allows integration, or compilation, of processes that in turn reduces the demands on working memory that occur in the performance of domain-relevant tasks compared with other tasks. This compilation hypothesis predicts that age-related deficits in a skilled domain should be smaller among people with more training than among those with less training (i.e., the answer to the first question should be *yes*). In addition, to the extent that encoding of poorly structured materials must be based on noncompiled processes, there should be larger age differences with poorly structured materials than with well-structured materials, at least among experienced people (so the answer to the second question should also be *yes*). However, it is not at all obvious whether age-related differences should differ qualitatively from experience effects (Question 3). Indeed, both older persons and unskilled individuals should show functional deficits in working memory capacity during task performance.

This article addresses all three questions raised above: It examines whether experience minimizes age differences in cognition, whether poorly structured materials increase age differences, and whether age effects are qualitatively similar to experience effects. We attempt to find answers to these questions in music cognition, a domain in which individuals differ greatly in training and application of skills.

Previous work has suggested that all music listeners must share some basic musical knowledge to make sense of what they are listening to. For example, most adults, even musical novices, have learned at least the more basic tonal principles of their culture and use their knowledge in music processing tasks, as revealed, for example, by effects of tonal structure on preference ratings of melodies (Cross, Howell, & West, 1983) and by differences in accuracy of perception and recognition of more and less tonal melodies (Francès, 1958/1988). Such tonality effects appear around 8 years of age and grow stronger thereafter at least to young adulthood (Andrews & Dowling, 1991; Zenatti, 1969). At the same time, other research has shown that more highly trained musicians have more complex and subtle knowledge of the tonal idiom of their culture and use this knowledge to enhance their perception and memory of music as compared with those persons with less musical training (Dowling, 1978; Francès, 1958/1988; Krumhansl & Shepard, 1979).

Prior research on music cognition sets the stage for examining whether musical experience and tonal structure are used to more or less advantage in perception and memory of melodic stimuli as people enter old age. To begin to answer this complex question, we needed a task that required rapid and precise analysis of musical materials with varying degrees of musical structure. Such a task had to be sensitive to inefficiencies or slowdowns in processing, and yet it had to also show effects of tonal structure and musical experience. The transposition recognition task appeared well suited to our needs. Transposition in music means starting a tune on a different note, and thus a different key, but keeping all other aspects of the tune the same, including the size of the intervals. In this task, a *standard* melody presented one or more times in one or more keys is followed by a *comparison* that starts on a different note than the standard(s). The task is to judge whether the comparison is an exact transposition of the standard or whether it differs in one or more musical intervals.

Transposition recognition is trivially easy if the melodies are well-known: "Turkey in the Straw" sounds like "Turkey in the Straw" to even the least musically trained person and irrespective of the key in which the tune is played. This equivalence is not due to fuzzy encoding; a distortion of the melody in which a few intervals are altered will be easily rejected as not the real thing. Nonetheless, with novel melodies the transposition recognition task is subjectively quite difficult and highly error prone (Dowling & Fujitani, 1971). Performance of participants with novel melodies often differs only slightly from chance if the comparison melodies that are not transpositions (i.e., the lures) share melodic contour with their respective standards (i.e., if they have the same patterns of ups and downs in pitch; see Dowling, 1978). Same-contour (SC) lures evoke more false alarms then different-contour (DC) lures, and the false-alarm rate for SC lures can approximate the hit rate for true transpositions. Hence, discrimination between transpositions and SC lures is much lower than that between transpositions and DC lures.

Transposition recognition is known to improve as a function of participants' musical training as well as the melody's tonal structure. This point was established by Francès (1958/1988, Experiment 9) and has been often replicated (Dowling & Harwood, 1986). Although intuition may suggest that these two factors interact—that is, that participants with greater musical training would show a greater advantage of highly tonal melodies over less tonal melodies—Francès obtained the pattern of two main effects with no apparent interaction. Indeed, tonality effects in other transposition recognition experiments (Bartlett & Dowling, 1980) as well as other music processing tasks (Cross et al., 1983; Dewar, Cuddy, & Mewhort, 1977) have often been found to be as strong among musical novices as among more musically trained people. The possibility exists that the most sophisticated musicians may be able to encode short melodies without the need of a tonal anchor, in which case tonality effects would be attenuated among this group.

However, effects of both tonality and experience are likely to depend on the discrimination task required of the participants. Discriminating identical (ID) transpositions from DC lures (i.e., lures that differ from transpositions in the sequence of ups and downs in pitch height) appears to be highly accurate regardless of tonality of the melodic stimuli or participants' training in music. For example, Dowling (1978) reported that area-under-memory-operating-characteristics (MOC) scores for ID–DC discrimination with tonal melodies averaged .81 for participants with less than 2 years of musical training and .84 for those with 2 years or more, which was a nonsignificant difference (1.0 = perfect discrimination). Moreover, performance at about this same level (.89) was found by Dowling and Fujitani (1971) in a similar study using atonal melodies. These and other findings suggest that contour is a global attribute of melodies that can be encoded independently of precise interval information and more or less irrespective of musical experience. Interestingly, recent research suggests that contour information is well represented in immediate memory tasks but is forgotten more rapidly over filled retention intervals than more precise information about musical intervals needed for ID–SC discrimination (DeWitt & Crowder, 1986; Dowling & Bartlett, 1981; Dowling, Kwak, & Andrews, 1995; Edworthy, 1985). One interpretation is that contour information is explicitly represented as a global attribute of melodies in working memory but not in long-term memory (unless the melodies are well-learned). If this is so, and if older people are impaired at working memory capacity (Salthouse, 1990), the high levels of contour memory shown by young adult listeners may not be matched by older listeners. To the extent that memory for precise intervals is not a global memory skill, age effects for intervals might even be attenuated compared with memory for contour.

In summary, prior research on transposition recognition suggests higher performance among more musically trained participants and with more highly tonal melodies. Several studies suggest that these two effects will be independent of each other, but there are reasons to expect interactions if the participants span a wide range of experience. In any case, both experience effects and tonality effects should be stronger with the measure of ID–SC discrimination, which requires more fine-grained analysis of intervals, than with that of ID–DC discrimination, which might reflect working memory processes.

This article focused on differences between young adult and older listeners in transposition recognition tasks. Assuming that age-related deficits would be found, we considered three possible outcomes deriving from the questions raised earlier. First, if age-related deficits in cognitive performance tend to be minimized among those more expert in a domain (Salthouse, 1990), then age-related deficits in transposition recognition should be reduced among more musically trained listeners as compared with musical novices. Second, if older individuals suffer disproportionately from the use of poorly structured materials that mismatch their knowledge and skills (Wingfield et al., 1985), then age-related deficits should be increased with atonal as compared with tonal melodies. Finally, if age-related deficits involve general cognitive processes as opposed to domain-specific processes that are linked to experience, then age and experience should affect different measures in different conditions. One possibility that fits prior evidence is that experience (and tonality) affects ID–SC discrimination more than ID–DC discrimination, whereas age will show the opposite pattern.

General Method

Because the methods were quite similar in Experiments 1 to 4, we describe them in this section first. Exceptions to these procedures are noted in the separate experiments.

Participants

Demographics and recruitment. Older participants in these experiments were ages 60 to 80 and lived in the Los Angeles, CA area (Experiments 1, 2, and 3) or, in Experiment 4, in Dallas, TX or Lewisburg, PA. The younger participants were ages 18 to 30 and were undergraduate or graduate students enrolled at the University of California, Los Angeles or Bucknell University in Lewisburg, PA. The less musically trained older participants were recruited from senior citizen centers and adult education classes. This pool occasionally yielded people qualifying as musicians by our criteria (see criteria for musical experience section below). In Los Angeles, the older musicians were recruited partly from advertisements placed in the newsletter sent to a music appreciation club but mostly from an advertisement placed in the newsletter of the local musicians' union. The older Dallas musicians were recruited by personal contact with staff at the University of Texas—Dallas.

Some younger participants volunteered in partial fulfillment of a course requirement; the others volunteered without pay. Some older participants were tested in their homes; all older participants traveling to campus received reimbursement for parking fees. Older nonmusicians in Los Angeles and older musicians in Dallas had been promised a $10 honorarium during recruitment for other psychology experiments, which we paid. Other older participants did not receive this fee.

As far as we could tell, the younger and older participants were roughly equivalent in their socioeconomic status and educational level, allowing for the fact that the younger students had not yet completed college. That is, most of the older participants had attended college, with a fair number having earned advanced degrees. Although some people in both groups had been born outside of the United States, they had all had extensive exposure to Western musical traditions. Members of both groups reported themselves to be in good health. We conducted a basic auditory screening on the older participants, which is described below.

Criteria for musical experience. All participants completed a musical experience questionnaire, in which they were queried about the extent, frequency, and recency with which they engaged in various musical activities. These included listening to music, singing in amateur or highly selective choirs, studying an instrument privately, playing in a band or orchestra, giving solo recitals, and having any professional or semiprofessional experience.

We faced certain trade-offs in trying to equate musical experience in our older and younger groups. The higher the researcher sets the criterion for *highly experienced,* the less likely one will find people of ages 18 and 19 fulfilling the criterion. However, with an older group, one is more likely to find people with training that occurred some years in the past. How comparable are 5 years of training in the recent past of a college student and 10 years of training completed 30 years previously for a 60-year-old?

We decided to use the following criteria, recognizing that a perfect classification scheme was unlikely. The basic unit of classification was years of private lessons on an instrument, including voice. The criteria recognize that singers in amateur choirs may or may not have had private lessons, whereas anyone proficient enough to play in an instrumental ensemble is very likely to have studied privately. We also elected to ignore recency of training in our classification scheme. The classification units were as follows:

1. Unusable: People who claimed they seldom listened to music. No one fit this criterion.
2. Least trained: fewer than 2 years of lessons, or 4 years in an amateur choir.
3. Moderately trained: from 2 to 8 years of lessons, at least 10 years in an amateur choir, or at least 5 years in an amateur instrumental ensemble.
4. Highly trained: at least 8 years of lessons, 5 years in a highly selective choir, or any professional experience.

Occasionally we needed to modify the criteria, as in the case of self-taught musicians, who could be classified as more highly trained than strictly defined *years of study* might indicate. In contrast, people technically fulfilling a criterion of years of study might be placed into a lower category if they insisted that the lessons were unsuccessful. Particularly informative in this regard were the few instances where the participant claimed that the music teacher had suggested that the lessons stop! Experiment 1 used three levels of classification. Experiments 2, 3, and 4 used only two levels of classification: least and highly trained.

Materials

All melodies were newly composed, were seven notes long, and included at least four different chromas (note names). The 24 tonal melodies were musically pleasant and conveyed a strong sense of musical key. The 24 atonal melodies in Experiments 1, 2, and 3 were transformations of the tonal melodies by using the following algorithm: Except for notes C and D, each note in the tonal melodies was mapped onto the note one semitone below for the atonal versions. For example, a tonal sequence of G–B–D–C–A–B–G (shown in Figure 1) would become F#–A#–D–C–G#–A#–F#. These sequences were musically incoherent in that they did not establish a single key center. Experiment 4 used a slightly different scheme for producing atonal melodies, which is described later.

To familiarize participants with the idea of transposition and to assure an adequate level of performance on the transposition task, we presented each sequence four times in the acquisition phase of a trial. We presented the sequence in the original key, then transposed up by five semitones to the key of the subdominant, then transposed down from the original by five semitones to the key of the dominant, and finally it was heard in the original key once again.

After a brief retention interval, which was filled or unfilled depending on the experiment, we presented the comparison sequence. Comparisons were transposed up or down by three or four semitones, randomly determined, so that the comparison sequence was in a key different from any acquisition sequences. Comparisons were of three different types, occurring equally often: ID sequences were exact transpositions of the original. SC sequences differed in the fifth and sixth notes from an exact transposition of the original. However, these changed notes still preserved the contour, or pattern of ups and downs, of the original, as well as its tonality. DC sequences also differed in the fifth and sixth notes from an exact transposition of the original, but the changes resulted in a contour different from the original, while again preserving its tonality. The changed notes were chosen such that the average interval size was approximately the same in all three categories. Examples of the various trials are shown in Figure 1.

Sequences were presented at a rate of 3 notes/s in Experiments 1 and 4 and 3.33 notes/s in Experiments 2 and 3 and were rhythmically organized as shown in Figure 1, with slight accents on the first, fourth, and seventh notes. A silent period of 1.67 s elapsed between the offset of one acquisition presentation and the onset of the next; a 5.67-s retention interval elapsed between the offset of the acquisition phase and the onset of the test sequence. The trials were separated by a 10-s response period.

Several precautions were taken to ensure that the task could not be performed by simply using absolute pitch differences between the sequences. As already noted, comparison sequences were presented in a key different than any heard in the acquisition sequence. In addition, the original sequence began on a different key on each trial. More pre-

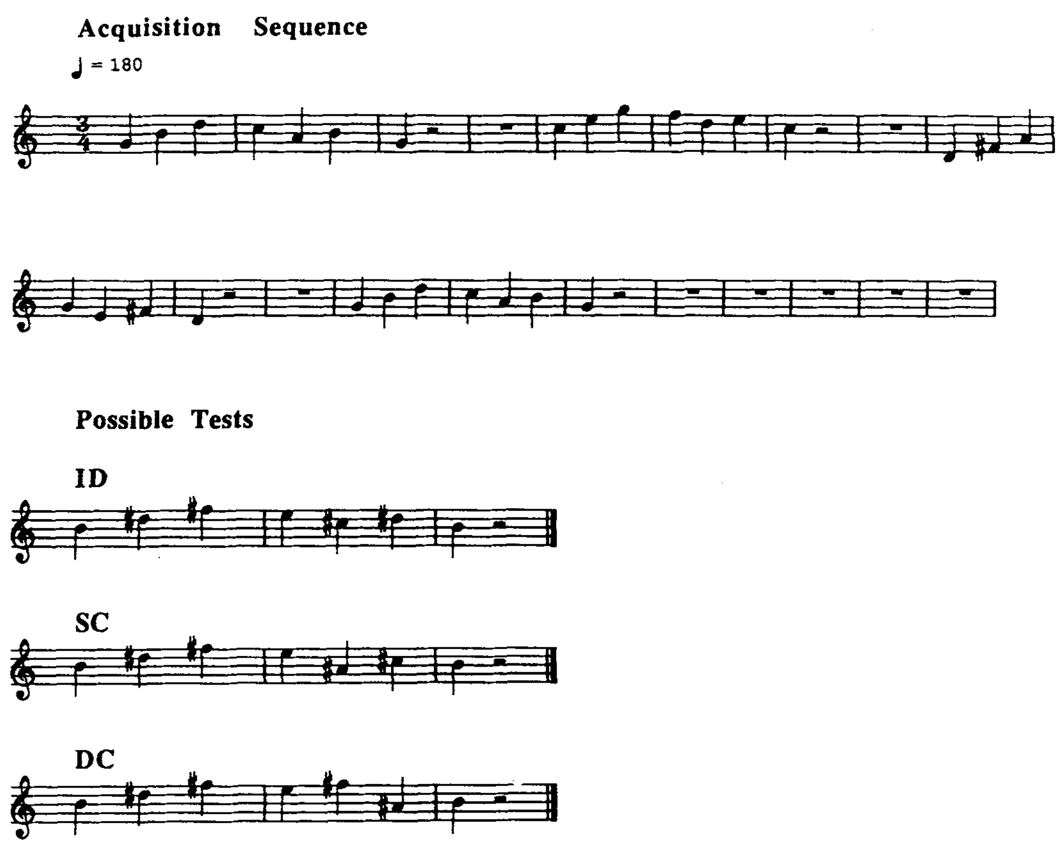

Figure 1. An example of the stimuli used in the transposition recognition task. ID = identical; SC = same contour; DC = different contour.

cisely, each of four blocks of 12 trials contained sequences in all 12 possible keys, in random order. Counterbalancing tapes were constructed for each experiment, such that every tune served equally often in an ID, SC, or DC sequence across participants. Approximately equal numbers of listeners performed the experiment with each tape. Sequence types were randomly intermixed on each tape.

Stimuli were produced on a Roland U-220 synthesizer using its "electronic piano" voice, under the control of Cakewalk software through a MIDI interface to a PC computer. Stimuli were recorded on audiocassettes for presentation over loudspeakers.

Procedure

An experimental session began with administration of the musical questionnaire. Because of specific recruitment and telephone contact, we were usually aware of a participant's classification before the session, but the questionnaire served to confirm (or sometimes change) that classification.

Older participants next received a brief audiometric screening. Using a Lucas GSI portable audiometer, thresholds were measured for pure tones ranging from 250 Hz to 8000 Hz. All participants had age-normal hearing in at least one ear. Some participants could not be screened because they wore a hearing aid. In those cases, we simply made sure that the volume of the music was sufficiently loud so that our participants could still take part in the experiment.

All participants then received the second half of the vocabulary test from the Wechsler Adult Intelligence Scale (WAIS). As each word was read aloud, participants wrote a synonym or brief definition on an answer sheet. The vocabulary test was used to provide a measure of cognitive performance that was dissimilar to the tasks under investigation and that was expected to show superior performance by the older participants. Thus, any age differences in the main task showing inferior performance among the older people would not be attributable to general cognitive impairments.

In Experiments 1, 3, and 4, the main transposition task was then introduced. In Experiment 2, another phase preceded the main task, as is explained in due course. All tapes were played on a high-quality stereo cassette player through the unit's speakers, and listeners were allowed to adjust the volume to their liking. To introduce the idea of transposition, the first practice trials used the familiar tune "Twinkle Twinkle Little Star." It was played four times and was transposed in the way described above. After 5 s, an ID, SC, or DC version of "Twinkle" was played, and participants were asked whether the comparison tune was an exact version of that tune. Participants used a 6-point answer scale: 1 = *very sure different*, 2 = *sure different*, 3 = *think different*, 4 = *think same*, 5 = *sure same*, and 6 = *very sure same*. The answer scale was in view at all times during the session. During the actual trials, all responses were written on an answer sheet by the participant, with the experimenter monitoring the correct placement of the answers in the blanks.

Practice trials eliciting incorrect answers were replayed and discussed by the experimenter. After participants successfully completed the trials with the familiar tune, another set of practice trials used an unfamiliar tune to demonstrate an ID, SC, and DC trial (the same unfamiliar tune was used for each example). On successful completion of the practice trials, the experimenter explained that the task would involve 49 similar trials, except that a different unfamiliar tune would be used on each trial. For Experiments 1 and 3, and in the atonal condition of Experiment 4, the experimenter added that some (all) of melodies would not be very tuneful (Experiment 2 used only tonal melodies).

The 48 experimental trials then ensued, preceded by an unscored warm-up trial. Participants were not permitted to hum, sing, play phantom pianos, or use any other external cues to aid their memories. If the 10-s response period was insufficient, the experimenter paused the tape to allow the participant to respond. A short rest period occurred about halfway through the 49 trials.

Sessions lasted between 45 and 75 min. Older participants were tested individually; younger participants were tested individually or in pairs.

Experiment 1

In the first experiment, we were interested in the effects of age, experience, tonality, and contour change on the ability to recognize transpositions. The first two variables were, of course, between-groups variables; the latter two were within-groups variables. The retention interval between acquisition sequence and comparison was unfilled. We predicted the obvious main effects: Younger would exceed older participants in this difficult short-term memory task, the more highly trained people would exceed the less highly trained, changes in tonal melodies would be more easily recognized than those in atonal melodies, and recognizing a difference would be easier in DC sequences compared with SC sequences.

One interaction of interest involved age and experience. If domain-specific experience allows the operation of strategies that may compensate for age-related memory difficulties, then we should see an Age × Experience interaction; the age difference should have been reduced for the more highly trained participants. We also hypothesized that older participants would benefit more than younger ones from the advantage of more structured materials, analogous to the findings in speech perception. Thus, we predicted an Age × Tonality interaction. Additionally, we thought that low and moderately trained people might benefit from tonality more than highly trained people, resulting in an Experience × Tonality interaction. Finally, we examined whether age would have its greatest effect when distinguishing ID from DC sequences was required and whether experience and tonality would have the greatest effects when distinguishing ID from SC sequences was required, following the reasoning outlined at the beginning of this article.

Method

Participants. Twenty-seven younger adults (18 women and 9 men; *M* age = 19.4 years, *SD* = 2.0) and 27 older adults (19 women and 8 men; *M* age = 69.3 years, *SD* = 5.5) participated in the study. Within each age group, 9 individuals were classified as least trained, 9 as moderately trained, and 9 as highly trained. All other details of participant recruitment and demographics were consistent with the description in the General Method section.

Materials. The main task used the three audio tapes described in the General Method section. One third of the participants in each of the six groups received each tape.

Procedure. The sessions proceeded as described above: musical background questionnaire, audiometry for the older participants, vocabulary test, practice task, and main task. The session took about 45 min for the younger participants and 1 hr for the older participants.

Results

All analyses were initially performed using the entire sample of 54 participants. However, it became apparent that one of the older moderately trained participants did not follow instructions, as scores in several conditions were below chance levels. We mainly report analyses using $N = 53$, but we note whether results changed substantially when this individual was included.

Vocabulary. The first measure we analyzed was vocabulary

Table 1
Mean Vocabulary Scores in Experiment 1

	Age group			
	Younger		Older	
Training level	*M*	*SD*	*M*	*SD*
Low	18.8	7.9	26.3	9.9
Moderate	20.2	6.7	26.6	5.0
High	22.8	7.8	29.9	4.4

Note. Maximum score = 40. Younger = 18–30 years; Older = 60–80 years.

scores. On the WAIS, each answer may receive a score of 0, 1, or 2 points. For the 20 items we presented, the maximum possible score was 40. This analysis included all 54 participants. Consistent with much previous work, the older participants (*M* score = 27.6) exceeded the younger participants (*M* score = 20.5). As shown in Table 1, there appears to be a trend toward better performance on the vocabulary test with increasing musical training, but an analysis of variance (ANOVA) on the vocabulary scores revealed only a significant effect of age, $F(1, 48) = 12.77$, $MSE = 51.80$, $p < .001$. Neither a main effect of experience nor an Age × Experience interaction was significant. Nevertheless, we report some analyses that use vocabulary scores as a covariate to examine whether important results might be modified by this variable.

Hits and false alarms. We present tables of hits and false alarms to examine whether any effects might have manifested themselves primarily as errors of omission or errors of commission. A *hit* is defined as saying *same* (using a 4, 5, or 6 on the answer scale) for an ID trial. A *false alarm* is defined as saying *same* for an SC or DC trial. Table 2 lists the proportion of hits and false alarms for each Age × Experience group, collapsed over the tonality variable. However, most of our discussion centers around the age and expertise groups separately, whose means are also noted in italics in the table.

Considering age differences first, we see that the younger participants had higher hit rates than the older participants; they also had lower false-alarm rates, but this superiority was confined to the DC sequences. Considering experience, we see only minimal differences in hit rates. More highly trained people did exceed novices in suppression of false alarms, and this was most prominent for SC items. These effects are examined more formally in the following analyses.

Area-under-MOC scores. The main dependent measures we report are areas under the MOCs for ID–SC and ID–DC discrimination (*area scores*). To compute the ID–SC and ID–DC discrimination scores, the confidence ratings made by each participant were used to derive hit rates for ID items and false-alarm rates for SC or DC items in each tonality condition at up to five criterion levels. The hit and false-alarm rates in each tonality condition defined a participant's MOC for that condition, and the areas under the curve were computed. The area under the MOC provides an unbiased estimate of proportion correct (Swets, 1973), varying from 1.0 (perfect discrimination) to .50 (chance).

The initial analysis of the area scores was an ANOVA containing four variables: age, experience, tonality, and contour (ID–SC vs. ID–DC scores). All four main effects were as predicted. Younger participants outperformed older participants (*M* score = .81 vs. *M* score = .73), $F(1, 47) = 5.68$, $MSE = .05$, $p < .03$, and performance rose from the least to moderately to highly trained groups (*M*s = .72, .78, and .81, respectively), $F(2, 47) = 3.99$, $MSE = .05$, $p < .05$. Tonal items were easier than atonal items ($M = .79$ vs. $M = .75$), $F(1, 47) = 7.05$, $MSE = .02$, $p < .02$, and ID–DC discriminations were easier than ID–SC discriminations ($M = .86$ vs. $M = .69$), $F(1, 47) = 108.10$, $MSE = .01$, $p < .001$.

Figure 2 depicts two 2-way interactions of interest (*M*s and

Table 2
Proportion of Hits and False Alarms for Age × Experience Groups for Each Trial Type in Experiment 1

	Experience level							
	Low		Moderate		High			
Trial type	Proportion	*SE*	Proportion	*SE*	Proportion	*SE*	*M*	*SE*
ID (hits)								
Younger	.81	.05	.84	.05	.87	.03	.84	.03
Older	.74	.07	.73	.05	.74	.06	.74	.03
M	.78	.04	.79	.04	.81	.04	—	
SC (FAs)								
Younger	.60	.04	.48	.07	.42	.07	.50	.04
Older	.53	.08	.42	.04	.42	.04	.46	.03
M	.57	.04	.45	.04	.42	.04	—	
DC (FAs)								
Younger	.10	.04	.09	.03	.11	.05	.10	.02
Older	.32	.07	.21	.08	.07	.04	.20	.04
M	.21	.05	.15	.04	.09	.03	—	

Note. ID = identical; SC = same contour; FA = false alarm; DC = different contour; Younger = 18–30 years; Older = 60–80 years.

*SE*s). Confirming our impression from the table of hits and false alarms, we found an Age × Contour interaction, $F(1, 47) = 4.13$, $MSE = .01$, $p < .05$, such that age differences were more extreme in the easier ID–DC discriminations than in the ID–SC discriminations, where they were in fact negligible. Also, tonality interacted with contour, $F(1, 47) = 33.80$, $MSE = .01$, $p < .005$. Tonal sequences were more successfully handled than atonal sequences but only in the ID–SC discriminations. In the ID–DC discriminations, tonality conferred no special advantage. Thus, we see a contrast in that age primarily affected ID–DC discriminations, whereas tonality primarily affected ID–SC discriminations.

Because contour interacted with both age and tonality, we conducted separate ANOVAs on the ID–SC and ID–DC scores. Results were very clear: For the ID–DC discriminations, the only significant effect was for age ($M = .91$ for younger vs. $M = .80$ for older), $F(1, 47) = 8.86$, $MSE = .034$, $p < .005$. Conversely, for ID–SC discriminations, the only significant effects were main effects for experience (*M*s = .62, .71, and .73 for ascending levels of training), $F(2, 47) = 3.85$, $MSE = .030$, $p < .03$, and tonality ($M = .74$ for tonal vs. $M = .63$ for atonal), $F(1, 47) = 24.30$, $MSE = .014$, $p < .001$.

We did not find three interactions of interest. Tonality and age did not interact, $F(1, 47) = 2.45$, $MSE = .016$, $p = .12$, as tonality conferred just as large a processing advantage for younger as older people. Similarly, tonality and experience did not interact, $F(2, 47) = .90$, $MSE = .016$, $p = .41$; tonality conferred as large an advantage to all experience groups. Finally, we also did not find a significant interaction between age and experience, $F(2, 47) = 0.31$, $MSE = .05$, $p = .74$ (though the age effect appeared to be slightly larger in the least trained group, *M* of the age difference = .11, than in the moderately and highly trained groups, *M*s of the age difference = .06 and .05, respectively). We are fairly confident that these failures to find interactions were not due to ceiling or floor effects. Means in the relevant cells ranged from .66 to .85, all within a range that could reasonably allow performance differences to emerge. Neither the four-way nor any three-way interactions were significant.

Vocabulary as a covariate. We reanalyzed the data entering vocabulary score as a covariate, because of the potentially confounding increase of vocabulary scores with experience. The main effect of the covariate was significant, $F(1, 46) = 6.95$, $MSE = .045$, $p < .02$. However, very few of our effects depended on experience, so we did not anticipate any large revision in our results with this analysis; and, indeed, this was the case. The only noteworthy change was that the main effect of experience was not significant in the analysis of covariance (ANCOVA), $F(2, 46) = 2.41$, $MSE = .045$, $p = .10$. Because the means were ordered in the expected direction for an experience effect, we tested the area scores in a contrast for linear trend. This test did reveal a significant linear trend for experience, $F(1, 46) = 4.22$, $MSE = .022$, $p < .05$.

Discussion

The results of this study gave us both some expected and some surprising outcomes. To consider some expected effects first, we had predicted the superiority of the younger people compared with the older people. This transposition task has several features that previous research has suggested would be sensitive to

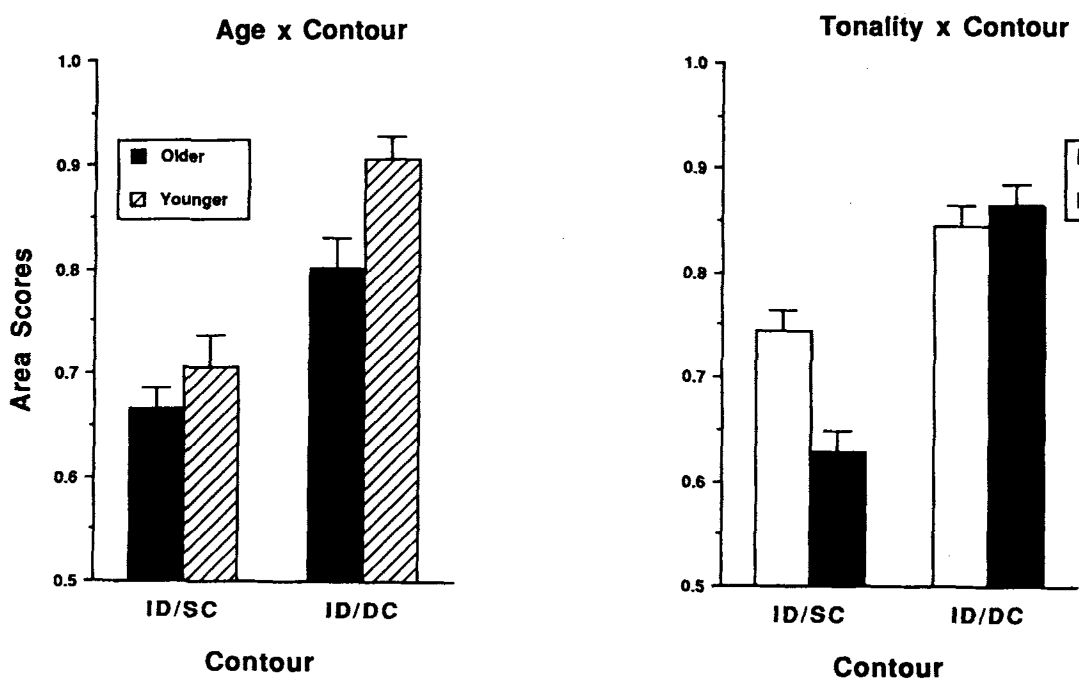

Figure 2. Age × Contour and Tonality × Contour interactions in Experiment 1. Error bars are standard errors. ID = identical; SC = same contour; DC = different contour; Older = 60–80 years; Younger = 18–30 years.

aging effects. The to-be-remembered item was novel, it was not easy to classify verbally, and the task required effortful abstraction and transformation of the music to detect the transposition successfully.

The superiority of the tonal materials over atonal materials was also unsurprising. Sequences with tonal centers offer several advantages in processing. First, they sound more familiar than atonal sequences, even to most musicians. Second, the establishment of a tonal center can serve as a reference point for all the intervals during both encoding and retrieval of the sequence. In atonal sequences, the successive intervals must be remembered one by one or perhaps in reference to the previous or following interval. Third, our tonal sequences conveyed particularly strong feelings of closure in their final interval. Because the changed notes were always in Positions 5 and 6 (out of 7) in the SC and DC sequences, their proximity to a very salient final reference interval might have been particularly helpful in our tonal materials.

The Tonality × Contour interaction is consistent with previous work (Dowling, 1991) and suggests how participants presumably carried out the task over short retention intervals. Recall that the tonality advantage occurred with the ID–SC discriminations but not the ID–DC discriminations. In the ID–DC discriminations, the changed notes changed the contour of the sequence. Thus, to answer correctly, the listener needed only to retrieve the pattern of ups and downs in the acquisition and test sequences. The identity of the intervals was irrelevant. In the ID–SC discrimination, contour could no longer be used to answer successfully because the acquisition and test sequence had the same contour. The listener had to retrieve the precise intervals from both sequences and maintain them for the comparison. Here, tonality provided the reference points useful in abstracting the interval information.

That the Tonality × Contour effect did not interact with either age or experience, or tonality with age or experience, suggests that all of our participants were using tonality in the same way and to the same degree. Taking age first, this suggests that both younger and older people are equally sensitive to the tonality manipulation and that both groups are capable of using it to aid memory. Considering experience, it is notable that untrained people pick up the constraints inherent in the tonal sequences, despite the lack of explicit tuition in such matters. However, even the musicians benefited from the processing advantage of tonality. We had thought that highly trained musicians might have developed other strategies for abstracting and using interval information that would supersede tonality. Alternatively, we had supposed that musicians would find the task so easy that tonality would not be needed in order to do well. Apparently neither assumption is true. However, we must remember that our highly trained group was not on the whole composed of world-class performers or musicians trained extensively in atonal idioms. Perhaps our original predictions would hold true for such groups.

The result we found most surprising involved effects of experience. Although the experience main effect was statistically significant, it was not large in size and disappeared entirely when the effects of vocabulary performance were controlled. Furthermore, experience did not reliably qualify the main effect of age. However, before we make the counterintuitive assertion that musical training is not a very powerful influence in musical tasks, we examine evidence from later experiments, in which we decreased the training levels to two while increasing sample size from 9 to 12 in each cell in most cases. Moderately trained participants generally performed at an intermediate level between the other two groups, so we feel that results do not suffer in clarity by this modification.

Perhaps our most intriguing results concerned the dissociation of the variables affecting the ID–SC compared with the ID–DC discriminations, as seen in the Age × Contour interaction (and the separate analyses of the ID–SC and ID–DC discriminations), as well as in the hit and false-alarm rates of SC and DC items. These patterns suggest that the ability to abstract and compare contour in the ID–DC discrimination may be a general perceptual strategy that is vulnerable to decline with age. Success in the ID–SC discriminations depended more on a domain-specific ability, and thus was influenced by one's musical training. Consistent with domain specificity, this more musical task was the one influenced by our musical factor of tonality, but was less sensitive to the general factor of age.

Experiment 2

We had several purposes in Experiment 2. One goal was to replicate the results of Experiment 1 with a new, larger sample of participants in each cell. Another was to investigate the effect of a previous presentation of some of the items on our main transposition recognition task (priming). Finally, we were interested in how well each age and experience group would perform on a recognition test for those previously presented items, compared with the pattern on the transposition task. Although our interest in priming was a major motivation when we originally designed this study, this manipulation turned out to have no reliable effect on performance of our transposition task. Thus, we do not discuss the priming variable extensively in this article.

In summary, then, Experiment 2 was similar to Experiment 1, with the following exceptions: Half the items were presented in an initial phase for rating on pleasantness. During the main task, participants were asked for two judgments after each trial: an old–new recognition judgment followed by the transposition recognition. Because of the time demands of the added phase in this study, we eliminated the variable of tonality and only presented tonal items. As noted above, we also used only two experience groups with low and high levels of musical training, respectively.

With regard to transposition recognition, we expected to replicate the main effects of age, experience, and contour from Experiment 1. We also predicted that, once again, age would primarily affect ID–DC discriminations and experience would primarily affect ID–SC discriminations. We were also interested in whether an Age × Experience interaction would be reliable in this study, despite the lack thereof in Experiment 1.

For the old–new recognition judgment, we predicted an age effect. We were less confident about finding experience effects. In a pilot study, we gave older and younger participants pairs of tunes differing by one note in a same–different discrimination task, followed by an old–new recognition test for those items. The melodies were all tonal and contained six notes, making them similar to the current materials. After the discrimination task, all melodies were presented in an old–new recognition paradigm. Although the discrimination task differed in several ways from the current transposition task, we did find both age and experience effects in that task. In old–new recognition performance, the younger people exceeded the older people in memory performance, but musicians, defined post hoc as having more than 5 years of musical training, did not exceed nonmusicians. On the basis of these data and those from Experiment 1, we were more confident of finding experience differences on our transposition task than on the old–new recognition task.

Method

Participants. Participants included 24 younger and 24 older people; 12 in each group were classified as having a low level of musical training and 12 a high level of training. The younger people (17 women and 7 men) had a mean age of 19.4 years (SD = 2.3); the older people (12 women, 12 men) had a mean age of 70.7 years (SD = 5.0).

Materials. The 24 tonal items from Experiment 1 were used plus an additional 24 tonal sequences composed and counterbalanced in the same manner as previously. One third of the trials constituted ID, one third constituted SC, and one third constituted DC comparisons, as before. Each melody served equally often in each type of trial over the course of the experiment. For the pleasantness rating task, 24 sequences were randomly selected and recorded as an initial portion of each tape. The sequences were presented just as they were in the acquisition phase of the main experiment: original sequence, transposition up five semitones, transposition down five semitones, and original sequence again. Over the course of the experiment, each item served equally often as a presented and a nonpresented trial in the main task. This resulted in a total of six tapes for counterbalancing purposes.

Procedure. The musical background questionnaire, audiometry, and vocabulary test were administered as before. For the pleasantness rating task, participants heard the 24 tonal items, each presented four times in three different keys. Participants rated pleasantness of the tune on a 1 to 7 scale at any time during the four repetitions. No mention of a memory task was made at this point.

After the rating task, the transposition task was introduced and practiced as before. After the practice trials, participants were told that some items in the forthcoming transposition task may have appeared in the rating task. They were asked to indicate *yes* or *no* according to whether they believed the melody was previously presented. The old–new recognition judgment was made at any time prior to the completion of the acquisition part of the trial. That is, the judgment was made before the silent retention interval began. After the silent interval, the test sequence was presented, and participants then made their transposition judgment as in Experiment 1, using the 6-point scale. Sessions lasted about 1 hr for younger participants and 1 hr 15 min for older participants.

Results

Vocabulary. The mean scores (maximum = 40) for younger people with low and high training, followed by the same for the older people, were 20.1, 25.1, 28.3, and 30.4, respectively. Once again, older participants exceeded younger ones, $F(1, 44) = 11.95$, $MSE = 45.80$, $p < .001$, but the experience effect was not conventionally significant, $F(1, 44) = 3.37$, $MSE = 48.80$, $p = .07$. However, in light of a possible weak relationship between vocabulary and experience, we once again report some analyses that use vocabulary scores as a covariate.

Priming. As noted previously, having presented an item earlier for ratings on pleasantness did not affect performance on the transposition task (M area scores = .85 for primed vs. .81 for unprimed items), and this factor did not interact even marginally with others in the experiment. This manipulation is not discussed further in this article.

Transposition recognition: Hits and false alarms. Table 3 lists the hits and false alarms in transposition recognition for each Age × Experience group, collapsed over the priming variable. Once again, mean values for age and experience groups are highlighted.

When age differences are considered first, a pattern very much like that in Experiment 1 emerges. The younger exceeded the older participants in hit rates and in false alarms to DC items. No age difference was apparent for SC items. Unlike Experiment 1, experience effects were apparent in all three item types: hits, false alarms to SC items, and false alarms to DC items. A formal analysis of performance in transposition recognition follows.

Area scores. Our main analysis used area scores. As in Experiment 1, we found a main effect of contour (M = .83 for ID–DC discriminations vs. .74 for ID–SC discriminations), $F(1, 44) = 48.60$, $MSE = .010$, $p < .001$. Contour did not interact with other factors, although the means suggest a Contour × Age pattern similar to that in Experiment 1: Younger people exceeded older people by .07 for the ID–SC comparison (.77 vs. .70) and by .11 for the ID–DC comparison (.89 vs. .78). The left panel of Figure 3 shows mean area scores (plus standard errors) for Age × Experience groups, collapsing over contour and whether the item had been presented earlier in the session or not. As is obvious from the figure, age had a main effect, $F(1, 44) = 8.27$, $MSE = .047$, $p < .01$, as did experience, $F(1, 44) = 38.50$, $MSE = .047$, $p < .01$. The Age × Experience interaction was significant as well, $F(1, 44) = 5.56$, $MSE = .047$, $p < .03$. All three effects remained significant when the analysis was repeated using vocabulary score as a covariate: The covariate itself gave $F(1, 43) = 7.28$, $MSE = .041$, $p < .01$. Subsequent analyses on the two separate experience groups confirmed that the age advantage for younger people occurred only among the

Table 3
Proportion of Hits and False Alarms for Age × Experience Groups for Each Trial Type in Experiment 2

	Experience level					
	Low		High			
Trial type	Proportion	*SE*	Proportion	*SE*	*M*	*SE*
ID (hits)						
Younger	.77	.06	.85	.03	.81	.03
Older	.56	.06	.85	.05	.71	.05
M	.67	.04	.85	.02	—	
SC (FAs)						
Younger	.45	.04	.27	.06	.36	.04
Older	.49	.05	.27	.06	.38	.05
M	.47	.03	.27	.04	—	
DC (FAs)						
Younger	.22	.07	.06	.02	.14	.04
Older	.31	.05	.10	.03	.21	.04
M	.27	.04	.08	.02	—	

Note. ID = identical; SC = same contour; FA = false alarm; DC = different contour; Younger = 18–30 years; Older = 60–80 years.

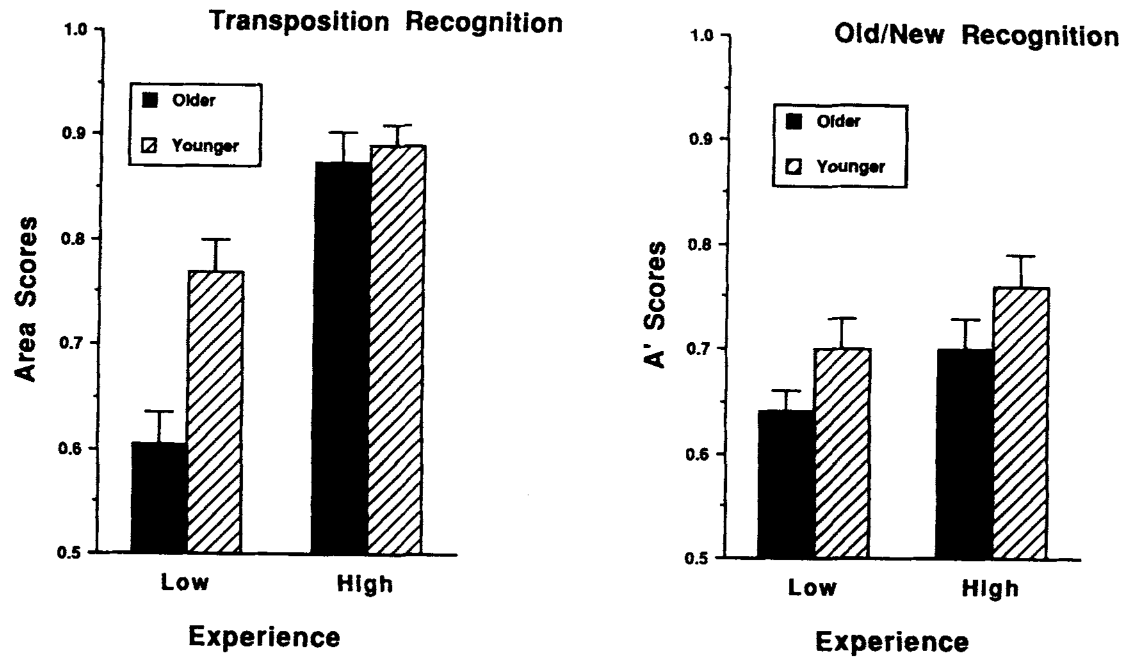

Figure 3. Experience and age patterns in transposition recognition and old–new recognition in Experiment 2. Error bars are standard errors. Older = 60–80 years; Younger = 18–30 years.

less trained, $F(1, 22) = 11.30$, $MSE = .057$, $p < .003$, and disappeared among the highly trained ($F < 1$).

Old–new recognition. We analyzed old–new recognition performance by computing the proportion of old judgments to old melodies (hit rate) and the proportion of old judgments to new melodies (FA rate) for each participant and then deriving an A' score (Grier, 1971). A' scores are estimates of area scores derived from single points on MOC plots (which was all we had for old–new recognition). As the right panel of Figure 3 shows, younger people exceeded older people, $F(1, 44) = 4.50$, $MSE = .010$, $p < .04$. There was a significant experience effect $F(1, 44) = 4.06$, $MSE = .010$, $p = .05$, but this effect disappeared when vocabulary score was used as a covariate. Age and experience did not interact; we certainly found no evidence for smaller age differences among more experienced people.

Given the contrasting patterns of area scores for transposition recognition and A' scores for old–new recognition, it seemed useful to conduct an ANOVA treating task as a within-subjects variable (not including vocabulary as a covariate and collapsing the ID–SC and ID–DC scores from transposition recognition). The ANOVA confirmed what Figure 3 suggests. There were main effects of age, $F(1, 44) = 8.76$, $MSE = .016$, $p < .005$; experience, $F(1, 44) = 24.30$, $MSE = .016$, $p < .001$; and task, $F(1, 44) = 27.50$, $MSE = .006$, $p < .001$. Of greatest importance, there was a significant Task × Experience interaction, $F(1, 44) = 17.80$, $MSE = .006$, $p < .001$, as well as a Task × Experience × Age interaction, $F(1, 44) = 4.83$, $MSE = .006$, $p < .05$. The interactions support our observations that, although the age difference in transposition recognition disappeared among the experts, the age difference in old–new recognition was maintained among experts.

Discussion

Several major findings from Experiment 1 were replicated in Experiment 2. In transposition recognition, we again found main effects of age and contour: Older people had more trouble than younger people, and changing the contour made recognition of a changed interval easier than when the contour was maintained. Similar to Experiment 1, this contour change advantage was the same in both experience groups. We continued to find a significant advantage of musicians over nonmusicians that was larger in magnitude here than in Experiment 1.

One difference from Experiment 1 was that, in Experiment 2, we found a significant Age × Experience interaction of which there were only hints in the prior study: Age differences among low experience participants were eliminated among the high experience participants. Although Experiment 1 showed a somewhat similar pattern, it was very weak and did not approach statistical significance.

Another difference between the two studies was the absence here of an Age × Contour interaction. Compared with Experiment 1, the ID–SC discriminations were a bit easier for both young and old, as might be expected once the difficult atonal ID–SC discriminations were eliminated. However, this difference between the studies should not be overemphasized. A comparison of Tables 2 and 3 shows a similar pattern between the experiments when considering hit and false-alarm rates. Specifically, both studies showed only minimal age differences in false-alarm rates to the difficult SC items, whereas age differences were more apparent in false-alarm rates to easier DC items and in hit rates to ID items. This pattern corresponds to the suggestion made in our discussion of Experiment 1 that tasks requiring access to general perceptual strategies (contour discrimination in our case) are more vulnerable to age-related decline than tasks requiring access to domain-specific skills (interval discrimination).

We attempted to avoid a strain on attentional resources resulting from dealing with two tasks, particularly in the older group, by having participants complete the old–new recognition task before making the transposition judgment. The overall level of performance was similar in Experiments 1 and 2 for both older and younger participants; grand mean area scores were within .02 across experiments for both younger and older groups. This close numerical performance appears to have validated our attempt. Despite this equivalence in performance accuracy, however, some differences in strategy may have been encouraged by the interfering effects of having to complete two tasks in Experiment 2. Another difference in method between the experiments was the presentation of mixed tonality lists in Experiment 1 and only tonal items in Experiment 2. Could either or both of these changes account for the two differences in results between Experiments 1 and 2? The next two experiments examine these possibilities. In the General Discussion section, we also consider the magnitude of the various effects discussed in terms of proportions of variance accounted for by each of the variables.

In contrast to the transposition test, the old–new recognition test showed only main effects of age and experience (weakly), without an interaction of the two. That is, musical training conferred no special advantage to older people in this task that required an explicit memory judgment and that involved a longer retention interval with more interference compared with the transposition task. These tasks were comparable on several dimensions: They were performed by the same participants, and answers in both tasks were given to an item at about the same point in the presentation sequence. Yet we found an Age × Experience interaction in the transposition test but not in the old–new recognition test. The advantage of experience in the old–new recognition task was quite minimal as shown by the lack of an experience effect in the analysis of covariance (ANCOVA). These findings, along with those of our pilot test, suggest that the transposition task is helped by domain-specific skills to a greater extent than is a memory test, even when that memory test consists of items within the domain of experience.

Experiment 3

In Experiment 3, we consider one variation in our methodology that might help clarify some of the differences in results between Experiments 1 and 2. In Experiment 2, participants had to complete an old–new recognition task before performing the transposition task on each trial. Although performance on transposition did not suffer in accuracy by adding the other task, perhaps the interference generated by having to make an explicit memory decision favored people who could call upon musical experience to aid in transposition recognition. Here, we introduced interference during the transposition task even more directly than in Experiment 2.

Experiment 3 returns to the methodology of Experiment 1, with one exception. Instead of silence during the retention interval, interfering melodies were played. To the extent that this more difficult task would call on music-specific knowledge, we expected larger effects of experience reminiscent of Experiment 2, compared with the weaker effect in Experiment 1. We predicted that results that had been seen so far in both our experiments would once again, be replicated here, and we were interested in seeing whether our unstable Age × Experience interactions would be solidified in this demanding task.

Method

Participants. Participants included 24 younger and 24 older people: 12 in each group had a low level of musical training, and 12 in each group had a high level of training. The younger people (18 women and 6 men) had a mean age of 19.9 years (SD = 1.9); the older people (14 women, 10 men) had a mean age of 69.8 years (SD = 6.1).

Materials. The tonal and atonal items from Experiment 1 were used. However, instead of a silent retention interval between the last acquisition sequence and the test sequence, participants heard a familiar melody played as an interference sequence. A total of 12 familiar melodies was used over the 49 trials that included well-known tunes like "Pop Goes the Weasel" and "You Are My Sunshine." In all cases, the 5.67-s retention interval was sufficient to play the first phrase of the familiar tune. The interference sequences were played in a distinctly different timbre from the acquisition and test sequences and were also played in a different key from acquisition and test sequences on a given trial.

Procedure. All procedural details were identical to those in Experiment 1, except that training trials included the interfering melodies. Participants were informed that they should "listen to the familiar melodies because we might ask questions about them later," but that they need not respond to them during the transposition recognition task. We actually asked the first few participants we tested to recall the names of as many familiar tunes as they could, after the main task was completed. Level of recall was very low, so this phase was discontinued. However, it was obvious that participants were listening to the familiar melodies during the task, as they often reacted with looks of recognition or tried to comment or hum along (this was, of course, discouraged). Using familiar tunes as interference was fortuitous in that the familiarity made the task more pleasant and partially offset the frustration at having those tunes serve as effective interference.

Results

Vocabulary. Means out of a maximum of 40 were 22.4 and 19.5 for younger nonmusicians and musicians and 29.4 and 28.3 for older nonmusicians and musicians, respectively. Again, older people exceeded younger people, $F(1, 44) = 23.50$, $MSE = 31.90$, $p < .001$, but in this case experience was clearly unrelated to vocabulary performance. Nevertheless, we continue to report some analyses using vocabulary as a covariate for reasons that soon become obvious.

Hits and false alarms. Table 4 lists the hits and false alarms in transposition recognition for each Age × Experience group, collapsed over the tonality variable. Marginals for age and experience groups are again italicized. The pattern of age effects was similar to that in the other experiments: Older people were more prone to make false alarms to the DC items but were equal to the younger people in SC false-alarm rates. Younger people showed only a small advantage in hit rates in this experiment. Highly trained people exceeded novices in both hits and false-alarm rates, with experience being quite advantageous in minimizing false alarms to both SC and DC items.

Area scores. As in previous experiments, the main dependent measures were area scores for ID–SC and ID–DC discrim-

Table 4
Proportion of Hits and False Alarms for Age × Experience Groups for Each Trial Type in Experiment 3

	Experience level					
	Low		High			
Trial type	Proportion	*SE*	Proportion	*SE*	*M*	*SE*
ID (hits)						
Younger	.67	.06	.76	.05	.72	.03
Older	.68	.05	.70	.04	.69	.05
M	.68	.04	.73	.03		
SC (FAs)						
Younger	.62	.04	.32	.04	.47	.04
Older	.57	.05	.38	.05	.48	.05
M	.60	.03	.35	.03		
DC (FAs)						
Younger	.29	.05	.06	.02	.18	.04
Older	.34	.05	.13	.04	.24	.04
M	.32	.04	.10	.02		

Note. ID = identical; SC = same contour; FA = false alarm; DC = different contour; Younger = 18–30 years; Older = 60–80 years.

inations. Variables were as in Experiment 1: age, experience, contour, and tonality. Main effects were found for experience (M = .80 for musicians vs. .65 for nonmusicians), $F(1, 44)$ = 25.90, MSE = .044, $p < .001$; tonality (M = .75 for tonal items vs. .70 for atonal items), $F(1, 44)$ = 6.81, MSE = .021, $p < .02$; and contour (M = .80 for ID–DC vs. .65 for ID–SC), $F(1, 44)$ = 77.30, MSE = .013, $p < .001$. Although a small difference in means favored younger people (.75 vs. .71), for the first time in this series, we found no significant main effect for age or any interactions involving this variable. The means do not even intimate an Age × Experience interaction; in fact, the age difference was slightly larger among musicians (.06) than nonmusicians (.02). Once again, we rejected the possibility that floor or ceiling effects were obscuring an interaction, as cell means ranged from .64 to .84.

As in Experiment 1 (tonality was not varied in Experiment 2), tonality interacted with contour, $F(1, 44)$ = 11.30, MSE = .008, $p < .002$. The form of the interaction was the same as we found there: Tonal and atonal items produced nearly identical ID–DC scores (Ms = .80 vs. .79, respectively), but tonal items evoked higher scores than atonal items in ID–SC scores (Ms = .70 vs. .61, respectively). No other interactions were significant. In particular, experience and tonality did not interact (cell Ms ranged from .63 to .84).

Vocabulary as a covariate. As we considered the absence of age effects, we thought it prudent to include vocabulary scores in an ANCOVA. Because older people had superior vocabularies, verbal skill might have in some way compensated for age-related decrements, giving the spurious impression that age had no effect in this task. In fact, the ANCOVA showed no main effect for age, $F(1, 43)$ = 3.39, MSE = .044, $p < .08$, and age did not interact with any other variables.

Discussion

Adding the interfering melodies reduced area scores by an average of .04 from Experiment 1 to Experiment 3, so this task was more difficult, as expected. However, the main effects of tonality and contour and their interaction were the same here as in Experiment 1 (and as in Experiment 2 for the main effect of contour). Thus the variables relevant to the task, rather than the participants, were operating as they had previously.

The two most striking results were the absence of age effects and the large experience effect. Considering the former result, we might ask whether the elimination of age effects were due to floor effects for everyone or to the reduction of performance among young people compared with previous experiments. Floor effects do not seem a likely explanation. The task was quite sensitive to experience, showing that a different independent variable elicited a wide range of performance levels, and, although a few conditions were very difficult, the grand mean area score in Experiment 3 was .73, comfortably above random performance.

Consistent with the increased importance of experience on this task, performance among musicians declined very little from Experiment 1 to the interference-laden Experiment 3 (average of .01 decline), whereas nonmusicians' scores declined by .07. We thus infer that the same conditions that equate performance between different age groups exaggerate the premium placed on musical knowledge. Adding interference during the retention interval can be presumed to occupy working memory. Any listener who can do so will call on special strategies to maintain the memory representation of the to-be-remembered sequence over the filled interval. These strategies could include explicit or implicit naming of intervals, recognition of chord patterns, or comparison with other familiar sequences.

Conversely, we proposed that age differences in the previous experiments have been largely confined to tasks requiring contour discrimination. Recall that one constant in our data is that false alarms to SC items never show an age difference (see Tables 2, 3, and 4). Contour discrimination has been shown to be sensitive to increased interference and increased retention intervals, whereas interval discrimination seems to improve as retention interval and interference increase (e.g., DeWitt & Crowder, 1986; Dowling, 1991; Dowling & Bartlett, 1981). These previous studies implied that contour is more efficiently processed in working memory over the short term, whereas interval information may be less dependent on immediate processing. The equality of performance between age groups when working memory is occupied and the superiority of young people in contour discrimination when working memory is unoccupied suggest that younger people may use their working memory more efficiently than older people under normal circumstances. When circumstances are not normal (e.g., the presence of compelling interference), the general perceptual strategies the young people excel at fade in importance compared with specific domain-relevant strategies used by experienced people but not novices at all ages.

Experiment 4

We designed Experiment 4 to help explain two discrepant findings regarding interactions with age between Experiments 1 and 2 (recall that Experiment 3 showed no age effects). First, Experiment 2 failed to find the reliable Age × Contour interaction of Experiment 1 (smaller age differences in ID–SC than

ID–DC comparisons), although the patterns of means in the hits and false alarms were similar across experiments. Although we did not have a clear idea about why this particular result should be found in one experiment but not another, we thought it prudent to determine whether this interaction was replicable with a large sample of participants in a no-interference paradigm.

Second, Experiment 2 showed an Age × Experience interaction that was only hinted at in Experiment 1. Perhaps the pure list presentation of tonal sequences in Experiment 2 enabled older experts to call on musical strategies more easily than when items were randomly alternating between tonal and atonal. In addition, the presentation of only tonal items in Experiment 2, which are more similar to everyday musical experience than are atonal items, might have advantaged particularly the older musicians. Both aspects of the difference in presentation lists could have been the cause of diminished age differences among musicians in that experiment.

Experiment 4 was a replication of Experiment 1, with one major and two minor modifications. The major modification was presenting the tonality variable as a between-subjects variable. Thus, half the participants received only tonal items and half received only atonal items. This allowed us to present items in a pure list, as in Experiment 2, but also allowed us to compare tonal and atonal items, as in Experiment 1. The between-subjects design also had the advantage of a replication using many more participants than our other experiments to check the stability of some of our previous findings.

We made one minor change by using two levels of musical training instead of three. The other minor change we made was using a stronger manipulation of tonality than we used in Experiments 1 and 3. In those experiments, atonal items were generated by modifying tonal items all in the same way. This resulted in items having a certain similarity among themselves and, with repeated exposure, may have allowed some participants to form a schema of sorts for the atonal scale. With a between-subjects design, the participants who heard only atonal items would have had even a greater chance of implicitly discovering the relationships inherent in our atonal pseudoscale. Thus, here we used more than one scheme for generating atonal items, which we hoped would make those items even less tonally coherent than were our previous items. This change allows a stronger test of whether tonality confers processing advantages differentially in our older or less trained participants. Other specific hypotheses included the replication of our four main effects: superiority of younger over older and trained over untrained participants and superiority of tonal over atonal items and ID–DC over ID–SC comparisons.

Method

Participants. A total of 89 people (48 younger and 41 older) participated in this study. The younger adults (27 women and 21 men) had a mean age of 19.2 years (SD = 1.3), and the older adults (28 women and 13 men) had a mean age of 70.4 years (SD = 5.8). Two musicians exceeded the age of 80. Among the younger people, 24 individuals were classified as being highly trained and 24 untrained; among the older people, 18 were highly trained and 23 were untrained. Half of each experience group received tonal items and half received atonal items.

Materials. The 48 tonal items were those used in Experiment 2. The 48 atonal items were composed in the following way. Each atonal item was based on an atonal scale that differed from the major and minor scales of western European music as much as possible. Furthermore, each was based as much as possible on a different atonal scale, so that participants could not become accustomed to the invariant underlying pattern of intervals in the atonal items. To achieve this, we constructed four interval patterns that spanned an octave with seven different pitch classes and that diverged as much as possible from the major and minor scales. Those four patterns had the following interval patterns (in semitones) between successive pitches: [3, 3, 1, 1, 1, 2, 1], [3, 1, 1, 2, 2, 2, 1], [3, 1, 2, 1, 2, 1, 2], and [3, 2, 1, 2, 1, 2, 1]. (The major scale pattern, for example, was [2, 2, 1, 2, 2, 2, 1].) We generated 28 atonal patterns by taking each of the seven pitch classes as the point of origin or *tonic.* Atonal items were randomly assigned to the pitch patterns, with each pitch pattern being used, at most, twice in the session. Six tapes were used; each sequence served equally often in an ID, SC, and DC comparison for the tonal and atonal conditions, respectively.

Procedure. The sessions proceeded as in Experiment 1, except that a given participant received either tonal or atonal items. The latter group was told that the sequences would not be very tuneful. Technical difficulties prevented administration of a hearing test in this experiment; however, participants were allowed to adjust volume of the cassette player to their liking.

Results

Vocabulary. The scores for younger people with low and high training followed by the same for older people were as follows: 22.9, 24.7, 24.6, and 32.1, respectively. As usual, older participants outperformed younger ones, $F(1, 81) = 27.50$, $MSE = .006$, $p < .001$, and experts exceeded novices, $F(1, 81) = 27.50$, $MSE = .006$, $p < .001$. Age and experience interacted, $F(1, 81) = 3.37$, $MSE = .006$, $p < .01$. As can be seen from the means, the experience effect is most notable among the older participants. To guard against the confounding of vocabulary score with experience, an ANCOVA using vocabulary score as the covariate was performed on the area scores. The main effect of the covariate was not reliable, and the analysis did not significantly modify any of the results from the ANOVA on area scores. Thus, we report analyses without vocabulary as a covariate in the sections to follow.

Hits and false alarms. Table 5 shows the proportions of hits and false alarms for age and experience groups in the same manner as previous experiments. As in all the experiments, the younger exceeded the older participants in hit rates, but no age difference was apparent in false alarms to SC items. Unlike the first three experiments, both age groups had similarly low false-alarm rates to DC items. As in all the experiments, musicians outperformed nonmusicians on all three types of items.

Area scores. Our main analysis once again used area scores. As in previous experiments, we found main effects of age (M = .83 for younger vs. .74 for older adults), $F(1, 81) = 25.00$, $MSE = .017$, $p < .001$; experience (M = .82 for musicians vs. .74 for nonmusicians), $F(1, 81) = 17.00$, $MSE = .017$, $p < .001$; and contour (M = .89 for ID–DC discriminations vs. .67 for ID–SC discriminations), $F(1, 81) = 420.90$, $MSE = .005$, $p < .001$. As in Experiments 1 and 3 (tonality was not varied in Experiment 2), performance on tonal items (M = .81) exceeded performance on atonal items (M = .76), $F(1, 81) = 5.52$, $MSE = .017$, $p < .05$. Replicating results from Experiments 1 and 3,

Table 5
Proportion of Hits and False Alarms for Age × Experience Groups for Each Trial Type in Experiment 4

	Experience level					
	Low		High			
Trial type	Proportion	*SE*	Proportion	*SE*	*M*	*SE*
ID (hits)						
Younger	.84	.03	.91	.02	.88	.02
Older	.66	.04	.77	.05	.72	.03
M	.75	.03	.84	.02	—	
SC (FAs)						
Younger	.55	.03	.50	.03	.53	.02
Older	.59	.04	.44	.05	.52	.03
M	.57	.02	.47	.02	—	
DC (FAs)						
Younger	.14	.03	.08	.02	.11	.02
Older	.19	.03	.06	.02	.13	.02
M	.17	.02	.07	.01	—	

Note. ID = identical; SC = same contour; FA = false alarm; DC = different contour; Younger = 18–30 years; Older = 60–80 years.

tonality also did not interact with either age or, in a simple way, with experience (but see next paragraph). Tonality did interact with contour in the same pattern as in Experiments 1 and 3: Tonality conferred no advantage in ID–DC discriminations (*M*s = .90 and .88 for tonal and atonal items, respectively), but did so in ID–SC discriminations (*M*s = .71 and .64 for tonal and atonal items, respectively), $F(1, 81) = 8.39$, $MSE = .005$, $p < .01$.

Unlike Experiment 1, age and contour did not interact (age effect = .09 and .10 for ID–SC and ID–DC items, respectively), and although tonality and expertise did not interact, we found a nearly significant three-way interaction of tonality, contour, and experience, $F(1, 81) = 3.78$, $MSE = .005$, $p = .06$. For completeness, we show the means of the Tonality × Contour × Experience interaction in Table 6, but the main importance of these interaction patterns will be considered in the General Discussion section when we discuss the magnitude of age and experience effects across experiments. Experience did not reliably modify age effects, although means hinted at such a relationship (age effect = .14 among the nonmusicians and .07 among the musicians; $p = .09$ for the interaction). Once again, we rejected the possibility that floor or ceiling effects were obscuring an interaction, as cell means ranged from .67 to .85. No other interactions were significant.

Discussion

Recall that this experiment was essentially a replication of Experiment 1, with a stronger manipulation of tonality that was varied between participants. These two changes did not appreciably change overall performance levels; means from the two experiments are in close agreement.

Adding Experiment 4 to our corpus allows us to be increasingly confident in some of our effects. We consistently found main effects of our experience and structural (i.e., contour and tonality) variables. No experiment yielded evidence that older people take more advantage of more musically coherent (tonal) materials than do younger people. Our two structural variables interacted in that tonality conferred advantages in abstracting intervals but not contour. Experienced people generally did not differ from inexperienced people in taking advantage of more structured material, with the exception of a hint in this last experiment that they may have abstracted intervals from tonal materials a bit more efficiently than nonexperts (see Table 6).

The Age × Experience interaction has an ambiguous status. Experiments 1 and 4 revealed patterns of means suggestive of the moderation of age effects with experience; this pattern was robust in the transposition task of Experiment 2. It was absent entirely in Experiment 3 and in our old–new recognition task of Experiment 2. In our view, this adds up to only weak evidence for a relationship between age and experience in the transposition task. A quantitative analysis of the strength of this and other effects will be considered in the General Discussion section.

One of our results from Experiment 1 was not replicated in Experiment 4, that of an Age × Contour interaction. The first experiment showed that age differences were apparent in easy ID–DC discriminations but not in hard ID–SC discriminations. Although subsequent experiments did not find this reliably in the ANOVA of area scores, we note that the tables of hits and false alarms do display a consistent pattern. In none of the experiments did we find age differences in false alarms to SC items (see Tables 2–5). We also note that no experiment found the reverse pattern of larger age differences in the harder (ID–SC) task, as might be expected if age differences increase as a simple function of task difficulty.

General Discussion

We conducted these studies to answer three questions about how age differences in cognitive performance are related to domain-specific knowledge and skills. The first question was whether age and experience interact such that age-related deficits in cognitive performance are reduced among persons who are experienced in a domain. The second question was whether age-related deficits in cognitive tasks are exacerbated when the stimuli are relatively unstructured and poorly matched to participants' knowledge and skills. The third question was whether age-related deficits reflect functional losses in expertise or perhaps losses in speed or efficiency of processing apart from actual knowledge.

Table 6
Area Scores For the Tonality × Contour × Experience Interaction in Experiment 4

	Musicians				Nonmusicians			
	Tonal		Atonal		Tonal		Atonal	
Comparison	Score	*SE*	Score	*SE*	Score	*SE*	Score	*SE*
ID–SC	.76	.02	.67	.02	.66	.02	.60	.02
ID–DC	.93	.02	.93	.01	.87	.02	.83	.02

Note. *SE* = standard error; ID–SC = identical transposition–same contour; DC = different contour.

We addressed these three questions in an underresearched domain, that of age-related differences in music cognition. We selected this domain because of the great range of musical knowledge and skill found within samples of educated adults. We focused on the task of melodic transposition recognition, because this task is sensitive to effects of experience as well as structure of materials, and it appeared to have properties that would make it age sensitive. Our findings confirmed effects of experience, musical structure, and age, and the findings spoke to each of the three questions we raised. We now discuss how our findings answered these questions and consider some of their broader implications for theories of aging and experience.

With respect to the first question of whether age differences are smaller among more highly trained than less highly trained people, the answer appears to be *sometimes,* as we noted in the previous section. To more fully consider just how the findings differed among experiments, it is useful to look beyond issues of statistical significance to examine the strength of the effects of our variables in terms of the variance that each accounted for. Table 7 (top half) displays ω^2 scores representing the proportions of the between-subject variance accounted for by the main effects of age and experience and by the Age × Experience interaction in all the experiments. It is clear that the main difference among the studies concerned not the Age × Experience interaction (which nowhere accounted for very much variance), but rather the pattern of main effects of age and experience. Whereas the age and experience effects had about the same strength in Experiments 1 and 4, the experience effect was much stronger than the age effect in Experiments 2 and 3.

How can this difference in outcome be explained? In an earlier discussion, we considered differences in outcome between Experiments 1 and 2. One difference in method between them was that Experiment 2 included a priming manipulation (half of the melodies had been previously presented), whereas Experiment 1 did not. Because all experiments included unprimed, tonal melodies, it is informative to examine the variance explained by age and experience and the Age × Experience interaction looking only at such melodies. As shown in the bottom half of Table 7, most effects appeared somewhat weaker when only unprimed–tonal melodies were considered than when all melodies were considered. Nonetheless, the essential pattern remains unchanged.

A second possible difference was that Experiment 2 included only tonal items compared with Experiment 1 that used a mixed list. However, Experiment 4 did not reveal an Age × Experience × Tonality interaction, suggesting that tonality condition did not influence the strength of any Age × Experience interaction. Thus, we conclude that neither the presence of primed items of Experiment 2 nor the presence of atonal items of Experiment 1 can account for the difference in the findings of these studies.

Third, participants in Experiment 2 were required to make an old–new judgment after each acquisition sequence and prior to its respective test sequence. The necessity of making this old–new judgment may have produced some amount of interference in memory for the melodies, making high-level musical encoding of the melodies a more advantageous strategy in Experiment 2 than in Experiment 1. A point that favors this argument is that the age and experience effects of Experiment 2 were quite similar to those of Experiment 3, which also included interference (albeit a different kind). This similarity is especially striking in the analysis of unprimed–tonal items common to the studies (Table 7).

To summarize, these findings question the notion that age-related deficits in cognitive performance are necessarily reduced among people experienced in a domain. Experiment 2 showed the largest effect of experience and a significant age effect, which perhaps made that experiment the most likely to show an interaction between those variables. Indeed, a significant interaction was found, but one accounting for only 5% of the variance. Although such interactions between age and experience are observed in some cases (Morrow et al., 1992), in the domains of chess (Charness, 1981) as well as bridge (Charness, 1983), researchers have found that age differences in some aspects of performance are actually increased among more highly skilled participants. Such findings contradict a "compilation" hypothesis discussed by Charness (1985, p. 253).

Although our data provided a rather mixed answer to our first research question, they spoke much more clearly to our second research question concerning whether age-related differences are increased when materials are poorly structured or when they mismatch participants' knowledge and skills (Poon, 1985). When applied to the domain of music cognition, this hypothesis carries the clear implication that age-related differences in detecting transposed melodies should be increased if the melodies are relatively atonal. We found no support for this implication in any of our studies, despite performance levels that would allow such an interaction to manifest itself. We therefore conclude that though there probably are cases where poorly structured stimuli increase the size of age differences (e.g., Wingfield et al., 1985), recognizing transposed melodies is not one of these cases.

We also found that experts do not as a rule exploit well-structured materials any more effectively than do novices in this domain. Aside from a nearly reliable Tonality × Contour × Experience interaction in Experiment 4, our tests showed no evidence for a Tonality × Experience interaction, again despite

Table 7
Variance Explained (ω^2) by Age, Experience, and the Age × Experience Interaction in Transposition Recognition

	Effect		
Item and experiment	Age	Experience	Age × Experience
All items			
Experiment 1	.08*	.08*	−.02
Experiment 2	.08*	.39*	.05*
Experiment 3	.01	.34*	−.01
Experiment 4	.18*	.12*	.02
Unprimed–tonal items			
Experiment 1	.11*	.06	−.03
Experiment 2	.02	.28*	.03
Experiment 3	.04	.29*	.01
Experiment 4	.15*	.11*	.01

Note. ω^2 can take on values less than 0 when the effect size is very small. These values should be considered as essentially 0.
* $p < .05$ by an F test.

appropriate performance levels. Other research has shown that people minimally trained in music, even among grade school students, are sensitive to the differences among melodies in which all notes come from within a standard musical scale and those in which some notes come from outside the scale (Cross et al., 1983; Krumhansl & Keil, 1982). To our knowledge, only a few studies (Cohen, Trehub, & Thorpe, 1989; Dowling, 1984) have shown that more experienced people take more advantage of musically well-structured materials than do less experienced people. Our research suggests that nonexperts can not only discriminate atonal from tonal sequences but can use that information to detect interval and contour changes across transposition. This ability remains robust until older age, as we never encountered interactions of age, experience, and tonality.

The third question we asked was whether age-related deficits mirror experience effects such that deficits among older persons as compared with young adults resemble deficits among those unskilled in a domain as compared with those more skilled. Our data provide strong indications that the answer to this question is *no*. In fact, three sets of findings suggest that age effects and experience effects are qualitatively different.

First, in Experiment 1, we found significant age differences in ID–DC discrimination, but not in ID–SC discrimination. Conversely, we found a significant experience effect in ID–SC but not ID–DC discrimination. Although this finding makes sense, two caveats are required. First, whereas Experiment 1 showed a significant Age × Contour interaction, it did not show a significant Experience × Contour interaction. We simply found that experience reliably affected ID–SC discrimination but not ID–DC discrimination. Second, neither the Age × Contour interaction nor the Experience × Contour interaction were significant in the remaining experiments. However, a perspective on these outcomes can be derived, once again, from measures of the variance explained by age and experience across the three experiments.

Table 8 displays ω^2 scores for variance explained by age, experience, and Age × Experience interactions, with each of our measures in all experiments. Considering Experiments 1 and 2, note that experience accounts for greater proportions of the variance in ID–SC discrimination than in ID–DC discrimination. Conversely, age accounts for greater proportions of variance in ID–DC than in ID–SC in those experiments (no age effects were found in Experiment 3).

Experiment 4 presents an interesting gloss on this pattern. Recall that we obtained a nearly significant interaction of contour, tonality, and experience. We therefore looked at the magnitude of age, experience, and Age × Experience effects in the four types of stimulus items separately (tonal and atonal items and ID–SC and ID–DC comparisons). Table 8 confirms that Age × Experience effects are negligible for all the comparisons. The tonal items (see the first two lines under Experiment 4 in Table 8) essentially constituted a replication of the tonal condition in Experiment 1 except that Experiment 4 presented a pure list of tonal items. Note that the pattern from Experiment 1 is replicated: Experience effects are larger than age effects for ID–SC items, but the reverse is true for ID–DC items. Earlier, we interpreted this pattern as suggesting that relatively age-invariant domain-specific skills are particularly important in interval processing, whereas age-related perceptual skills are more important in contour processing.

Table 8
Variance Explained (ω^2) by Age, Experience, and the Age × Experience Interaction in ID–SC and ID–DC Discrimination in All Experiments

Experiment and comparison type	Effect		
	Age	Experience	Age × Experience
Experiment 1			
ID–SC	.01	.10*	−.03
ID–DC	.13*	.03	.00
Experiment 2			
ID–SC	.04*	.41*	.05*
ID–DC	.10*	.29*	.03
Experiment 3			
ID–SC	−.01	.31*	−.01
ID–DC	.02	.26*	−.01
Experiment 4			
ID–SC (tonal)	.08*	.15*	.02
ID–DC (tonal)	.19*	.03	.00
ID–SC (atonal)	.20*	.07*	.01
ID–DC (atonal)	.15*	.13*	.01

Note. ID = identical; SC = same contour; DC = different contour. ω^2 can take on values less than 0 when the effect size is very small. These values should be considered as essentially 0.
* $p < .05$ by an F test.

The atonal items in Experiment 4 were more atonal than in the other experiments because of the use of four different pseudoscales to generate the items. In addition, presenting atonal items in a pure list makes instantiation of musical schematic knowledge especially difficult and forces listeners to use other perceptual strategies. Consistent with these intuitions, we see that for the first time, age effects become larger than experience effects in ID–SC comparisons. Age and experience seem to contribute equally to the overall variance for ID–DC comparisons with difficult atonal items. These differing patterns of age and experience are evidence in favor of the qualitative differences argument we make here.

A second piece of evidence for a qualitative difference between effects of age and experience came from Experiment 2. Recall that Experiment 2 showed an Age × Experience interaction in transposition recognition. However, it showed no such interaction in old–new recognition, as shown in Figure 3. Once again, an assessment of ω^2 scores provides an interesting perspective on the pattern of means. As mentioned earlier, Table 7 reveals that the main effect for experience was considerably stronger than the main effect for age or the Experience × Age interaction in the transposition recognition task (the ω^2 scores were .39, .08, and .05, respectively). In contrast, the experience main effect was comparably weak in old–new recognition (.06 for experience vs. .07 for age and −.02 [essentially 0] for the interaction). We conclude that experience affected performance much more strongly than did age, but only for transposition recognition, not for old–new recognition. The pattern makes sense if the effects of experience reflect a musical encoding strategy on the part of experienced participants, whereas the effects of age reflect general processing impairments affecting memory for melodies. Musical encoding would be expected to be useful for transposition recognition, but not necessarily for old–new recognition. In contrast, general memory processes may be important in both tasks.

A third source of evidence that age and experience have differing effects derives from the comparison of Experiments 1 and 3. The two studies were procedurally identical except that Experiment 3 included melodic interference between each standard melody and its comparison in the transposition recognition task. However, Experiment 1 produced effects of age and experience, whereas Experiment 3 produced only an experience effect. Moreover, the experience effect in Experiment 3 was much larger than that in Experiment 1 (see Table 7).

Again, the pattern might be explained by the hypothesis that experience effects reflect musical encoding whereas age-related differences reflect general processing impairments including those of working memory. We think that the interfering melodies of Experiment 3 minimized the role of working memory because the limited-capacity memory store was fully occupied in processing the interfering melody, even if later recall was not required. A more abstract encoding of chroma information may withstand better the effects of interference among more experienced participants because of the ability to link incoming information immediately to information such as interval relations available in a long-term memory store.

The same reasoning can account for the large role that experience played in Experiment 2. Making the old–new judgment between the standard and comparison tune could reasonably be considered a form of interference, once again putting a premium on knowledge-based mnemonic strategies. Indeed, experience once again accounted for an impressive amount of variance, whereas age effects were much smaller.

Finally, a fourth piece of evidence in favor of our argument comes from our analysis of hits and false alarms. Low experience consistently was associated with lower hit rates and higher false-alarm rates compared with high experience (see Tables 2–5). However, older people produced lower hit rates than younger people (they were nearly equivalent to younger people in Experiment 3) and usually produced higher false-alarm rates to DC items but never produced higher false-alarm rates to SC items. Clearly, the advantages of youth and experience are reflected in different patterns of performance.

Our findings regarding age and experience are consistent with the familiar idea (e.g., see Charness, 1985, p. 243) that age-related differences reflect damage or deterioration in biological "hardware," whereas experience effects reflect improvements in task strategies, or "software." The validity of this metaphor is important to determine, because it bears on the question of whether "use it or lose it" is useful advice or just an analgesic falsehood that in the long term only hurts. This clearly is a matter for future research that assesses brain function as well as behavior.

In summary, the effects of age and experience differed in that in the most musical versions of the task, age affected ID–DC discrimination more strongly than ID–SC discrimination, whereas experience affected ID–SC discrimination more strongly than ID–DC discrimination (Experiment 1). However, age effects were more pronounced in the least musical version of the task (atonal condition of Experiment 4). The two variables also differed in that the effects of experience exceeded those of age in transposition recognition, but not in old–new recognition of melodies (Experiment 2). Versions of the task that involved interference also showed strengthened effects of experience while weakening those of age. Finally, the patterns of hits and false alarms differed for age and experience.

We close by relating our work to some recent discussions in the literature about whether age-related declines reflect global deterioration of cognitive skills or, rather, can be attributed to inefficiencies in particular cognitive components of a task. If the latter is true, then age effects should be magnified as tasks increase in complexity but still require the same basic processing operations. Salthouse (1992) provided evidence that performance on a simpler task, plus a measure of working memory, was highly predictive of performance on more complex versions of the task for young as well as older adults. Age effects were exacerbated in these more complex tasks, presumably because more complex versions require more access to working memory resources that are less available with increasing age.

In contrast to the above findings, our age differences were *minimized in the harder* task. Why might this be so? Although our ID–SC discrimination was more difficult than the ID–DC discrimination, this was due to qualitatively different task requirements (coding of intervals vs. contour) as opposed to simply increasing the number of processing steps in the harder task. One possible explanation is that age is related to a processing resource, such as working memory capacity, that contributes to memory for contour of melodies but has little to do with detecting transpositions. In contrast, musical skill is related to scale-step or chroma-based encoding that benefits the process of detecting transpositions but is only weakly related to general memory processes such as contour detection or episodic recognition. Thus, age-related differences in transposition recognition do not reflect a regression in skill. They appear, instead, to involve domain-general processes such as working memory that exist and decline independently of skill. Our research, along with Salthouse's (1992), encourages the conclusion that both quantitative and qualitative differences among tasks may be important predictors of when age will and will not be associated with impaired performance.

References

Andrews, M., & Dowling, W. J. (1991). The development of interleaved melodies and control of auditory attention. *Music Perception, 8,* 349–368.

Bartlett, J. C., & Dowling, W. J. (1980). The recognition of transposed melodies: A key distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 501–515.

Charness, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology, 33,* 1–16.

Charness, N. (1981). Aging and skilled problem solving. *Journal of Experimental Psychology: General, 110,* 21–38.

Charness, N. (1983). Age, skill, and bridge bidding: A chronometric analysis. *Journal of Verbal Learning and Verbal Behavior, 22,* 406–416.

Charness, N. (1985). Aging and problem solving performance. In N. Charness (Ed.), *Aging and human performance* (pp. 225–259). New York: Wiley.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Cohen, A. J., Trehub, S. E., & Thorpe, L. A. (1989). Effects of uncertainty on melodic information processing. *Perception and Psychophysics, 46,* 18–28.

Cross, I., Howell, P., & West, R. (1983). Preferences for scale structure in melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 444–460.

Dewar, K. M., Cuddy, L. L., & Mewhort, D. J. K. (1977). Recognition memory for single tones with and without context. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 60–67.

DeWitt, L. A., & Crowder, R. (1986). Recognition of novel melodies after brief delays. *Music Perception, 3,* 259–274.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85,* 341–354.

Dowling, W. J. (1984). Musical experience and tonal scales in the recognition of octave-scrambled melodies. *Psychomusicology, 4,* 13–32.

Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception and Psychophysics, 50,* 305–313.

Dowling, W. J., & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology, 1,* 30–49.

Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America, 49,* 524–531.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition.* New York: Academic Press.

Dowling, W. J., Kwak, S. Y., & Andrews, M. (1995). The time course of recognition of novel melodies. *Perception and Psychophysics, 57,* 136–149.

Edworthy, J. (1985). Melodic contour and musical structure. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 169–188). New York: Academic Press.

Francès, R. (1988). *The perception of music.* (W. J. Dowling, Trans.). Hillsdale, NJ: Erlbaum. (Original work published 1958).

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75,* 424–429.

Krumhansl, C. L., & Keil, F. C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory and Cognition, 10,* 243–251.

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

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

Poon, L. W. (1985). Differences in human memory with aging: Nature, causes, and clinical implications. In J. D. Birren & K. W. Shaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 427–462). New York: Van Nostrand Reinhold.

Salthouse, T. A. (1984). Effects of age and skill in typing. *Journal of Experimental Psychology: General, 113,* 345–371.

Salthouse, T. A. (1990). Cognitive competence and experience in aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 310–319). New York: Academic Press.

Salthouse, T. A. (1992). Why do adult age differences increase with task complexity? *Developmental Psychology, 28,* 905–918.

Salthouse, T. A., & Somberg, B. L. (1982). Skilled performance: Effects of adult age and experience on elementary processes. *Journal of Experimental Psychology: General, 111,* 176–207.

Smith, S. W., Rebok, G. W., Smith, W. R., Hall, S. E., & Alvin, M. (1983). Adult age differences in the use of story structure in delayed free recall. *Experimental Aging Research, 9,* 191–195.

Swets, J. A. (1973). The relative operating characteristic in psychology. *Science, 182,* 990–1000.

Wingfield, A., Poon, L. W., Lombardi, L., & Lowe, D. (1985). Speed of processing in normal aging: Effects of speech rate, linguistic structure, and processing time. *Journal of Gerontology, 40,* 579–585.

Zenatti, A. (1969). Le developpement génétique de la perception musicale [The genetic development of musical perception]. *Monographes Françaises de Psychologie, Whole no. 17.*

Received June 24, 1994
Revision received August 29, 1994
Accepted September 19, 1994 ■