

Auditory imagery from musical notation in expert musicians

WARREN BRODSKY and AVISHAI HENIK
Ben-Gurion University of the Negev, Beer-Sheva, Israel

BAT-SHEVA RUBINSTEIN
Tel Aviv University, Tel Aviv, Israel

and

MOSHE ZORMAN
Levinsky College of Education, Tel Aviv, Israel

Anecdotal evidence has suggested that musical notation can trigger auditory images. Expert musicians silently read scores containing well-known themes embedded into the notation of an embellished phrase and judged if a tune heard aloud thereafter was the original theme (i.e., *melodic target*) or not (i.e., *melodic lure*). Three experiments were conducted employing four score-reading conditions: normal nondistracted reading, concurrent rhythmic distraction, phonatory interference, and obstruction by auditory stimuli. The findings demonstrate that phonatory interference impaired recognition of original themes more than did the other conditions. We propose that *notational audiation* is the silent reading of musical notation resulting in auditory imagery. The research suggests that it also elicits kinesthetic-like phonatory processes.

The focus of empirical investigation of imagery has been on the visual modality. However, several recent initiatives have concerned auditory imagery (for extensive reviews, see Godoy & Jorgensen, 2001; Reisberg, 1992). In the form of inner speech, auditory images serve an important function of human cognition in tasks such as verbal memory (rehearsal), speech perception, silent reading, and thought processes. This type of imagery, also referred to as *subvocalization* or *inner ear*, is the experience of an inner voice without vocal output or environmental input (Intons-Peterson, 1992; Smith, Reisberg, & Wilson, 1992). The seemingly auditory quality of inner speech is not necessarily auditory in origin, and it should not be attributed exclusively to events or systems that are strictly auditory in nature. For example, Baddeley and Logie (1992) found that echoic memory (i.e., temporary storage of auditory material) operates only in the presence of auditory stimuli and, hence, cannot be the seat of auditory imagery. Additionally, Mackay (1992) found that inner speech is linked to the phonological system. Such findings may, then, negate the common assumption

that there exists a single seat or locus of auditory-based imagery.

A special case of auditory imagery is *musical imagery*. In a host of experiments, Hubbard and Stoeckig (1988, 1992), Halpern (1988, 1989, 1992, 2001), Halpern and Zatorre (1996), and Zatorre (1994) had subjects create an image of raising a tone or chord a full step, set metronome beats to imagined music, tap out the tempo of imagined songs, imagine a song's starting tone, judge the similarity between tones heard externally and those heard internally, and mentally scan a melody for pitch and duration characteristics of heard stimuli. Collectively, these researchers found that musical images are generated in real time, encode fairly precise information about tempo and pitch, and contain information concerning melodic and harmonic relationships. Hence, musical images possess a sensory quality that is similar to the experience of perceiving. Moreover, EEG studies (Beisteiner, Altenmüller, Lang, Lindinger, & Decke, 1994), PET studies (Halpern & Zatorre, 1999; Zatorre & Halpern, 1993; Zatorre, Halpern, Perry, Meyer, & Evans, 1996), and event-related potential (ERP) studies (Janata, 2001a, 2001b) have shown that imaginal and perceptual processes for melodies share neural structures and topographies. But, for the majority of musically naive (untrained) people, music may still be a predominantly external experience. For example, musically naive individuals have been found to be significantly worse than highly trained individuals on all auditory imagery tasks (involving both music and everyday sounds), but not on visual imagery tasks (Aleman, Nieuwenstein, Bocker, & de Haan, 2000). The ability of

The first author wishes to acknowledge the support of the Council for Higher Education (VaTat Committee), which granted the postdoctoral research fellowship under whose auspices this research took place. This research was supported in part by the Israel Science Foundation, founded by the Israel Academy of Sciences and Humanities. The authors are indebted to Bruno Repp and two anonymous reviewers for their comments on two earlier versions of the manuscript. Correspondence should be addressed to W. Brodsky, Music Science Research, Department of the Arts, Ben-Gurion University of the Negev, P. O. Box 653, Beer-Sheva 84105, Israel (e-mail: wbrodsky@bgumail.bgu.ac.il).

musicians to experience musical images, then, may be *the* outstanding mark of a musical mind. Musical images have been viewed as a necessary condition for learning, retention, recall, recognition, and the anticipation of musical events (Seashore, 1938).

The research literature is unequivocal in pointing to the underlying cognitive system, which accounts for the maintenance and processing of auditory representations as expressed in musical images (for a review, see Kalakoski, 2001). One question often raised is the following: To what extent do musical imagery and the phonological working memory overlap? This is especially interesting because vocal music has many acoustic features in common with vocal and subvocal speech. Given that interference of imagery has been found to be most effective when both the image and the distracting stimuli originate from the same sensory modality (Intons-Peterson, 1992), and given that auditory imagery has been blocked most effectively by auditory input (Smith et al., 1992), what, then, is to be expected in the case of musical imagery? Will musical imagery be affected most by the presentation of irrelevant musical, nonmusical auditory, or verbal material? The initial studies linked both auditory and musical imagery to the crisscrossing of aural and oral channels (Reisberg, 1992), and the triggering of musical images was also found to be linked to motor imagery (Mikumo, 1994; Petsche, von Stein, & Filz, 1996). More recently, PET studies (Halpern, 2001; Halpern & Zatorre, 1999) have found that the supplementary motor area (SMA) is activated during musical imagery, and most specifically during covert mental rehearsal (Langheim, Callicott, Mattay, Duyn, & Weinberger, 2002). Accordingly, the SMA may mediate rehearsal that involves motor processes such as humming. Furthermore, the role of the SMA during imagery of familiar melodies has been found to include both auditory components (of hearing the actual song) and carrier components (such as an image of subvocalizing, fingers moving on a keyboard, or someone else performing; Schneider & Godoy, 2001; Zatorre et al., 1996).

For some individuals, musical imagery can also be triggered by reading the graphic representation of music (i.e., musical notation). Following Gordon (1975, 1993), this phenomenon will be referred to here as *notational audiation*. Yet there are musicians who possess high-level skills without knowledge of how to read or write musical notation, and others can name all the notation symbols but cannot accurately audiate the sounds they represent. As a result, some researchers have reservations about notational audiation. For example, Sloboda (1974, 1984) claimed that musicians should be able to interpret a score before playing it, but that this should not be confused with the claim that they should also be able to hear what the music sounds like in their mind's ear. He defended his position as follows:

At first glance there seems to be an obvious falsifying counter-example to this null hypothesis. This is the case

of people who claim to be able to read a score in complete silence, without mediation of instrument or voice. However, I choose not to make anything of this example here. There are several reasons for this. One is that there is no obvious way of assessing the claim. Another is that among the population of fluent music readers those who can read silently are statistically rare, and so third, they may be employing processes of silent reading that are not normally employed in reading for performance (Sloboda, 1984, p. 224).

But musicians have always given credence to the presence of imagery from musical notation. Long ago, the Romantic composer Schumann (1848/1967) declared that his students must get to the point where they could hear the music from the page and be able to picture the piece at first hearing as though the score were in front of them.

To date, there is virtually no empirical evidence to support the notion of notational audiation. Most unfortunately, the two major research compilations on auditory imagery (Reisberg, 1992) and musical imagery (Godoy & Jorgensen, 2001) do not include so much as a research note on the subject. There is one music cognition study (Experiment 6 in Waters, Townsend, & Underwood, 1998) that shows that skilled musicians can successfully match one bar of musical notation to a subsequently presented auditory sequence in a *same-different* matching task. However, although the task is referred to as converting visual representations of music to auditory representations, the study does little to demonstrate that task performance is based on evoked musical imagery (rather than on structural harmonic analysis or visual surface cues found in the notation). In addition, an fMRI study (Nakada, Fujii, Suzuki, & Kwee, 1998) and a magnetoencephalography (MEG) study (Schurmann, Raij, Fujiki, & Hari, 2002) clearly show audiocortical activity while musical notation is read. However, these studies do little to reveal the cognitive process underlying this phenomenon. Therefore, as far as cognitive science is concerned, it appears that auditory imagery as a response to musical notation continues to remain essentially conjectural in nature.

The Present Study

The purpose of the present investigation was to develop an experimental paradigm to investigate notational audiation. To this end, the study exploits compositional techniques of theme and variation that allow a well-known theme to be visually disguised. Some notes are displaced in different octaves, some notes are heard for shorter durations, and new notes are added. Yet, always, the original melody, phrase structure, and harmonic plan remain intact as scaffolding (see Figure 1). The original well-known theme, then, is visually indiscernible, because it is embedded into a "new" embellished phrase. This resultant embellished phrase is referred to as an *embedded melody* (see Figure 2). Identification of the original well-known theme relies for the most part on aural

Figure 1. Embedded melody (i.e., variation, top stave) and original well-known theme (bottom stave) of “La donna è mobile” from the opera *Rigoletto* by G. Verdi (arrangement copyright by M. Zorman, 1997). The participants viewed only the embedded melody as it appears in Figure 2.

skills. On the basis of evidence (Deliège, 1996; Deliège & Melen, 1997; Melen & Deliège, 1995) that melody recognition depends on certain local surface cues abstracted during hearing (which are then memorized), it was assumed that embedded melodies would be useful in the empirical study of notational audiation.

Experimental task. The participants were required to silently read the score of an embedded melody and, then, on hearing a tune after the notation was removed from sight, to accept or reject the tune as the original well-known embedded theme. However, the ability to correctly recognize original well-known embedded themes might not necessarily be evidence in itself of notational audiation. There could be several other explanations of

how musicians are capable of successfully performing this task. For example, informed guesswork or harmonic analyses might be involved. Sloboda (1994) suggested that it is more than a possibility that trained musicians can compute what notation sounds like (at least to some degree of specification) just by examining a score, provided that they have a rich enough base of prior musical experience to fall back on. But if theme recognition could be hampered by distraction conditions (which may be assumed to engage audiation processes), then one could conclude that notational audiation was in operation during nondistracted score reading. Hence, several distraction conditions are employed in the study. It should be pointed out, however, that the use of distract-

Figure 2. Embedded melody employing the well-known theme “La donna è mobile” from the opera *Rigoletto* by G. Verdi (arrangement copyright by M. Zorman, 1997).

tion conditions allows for the further investigation of the auditory-versus-phonatory nature of imagery triggered by musical notation. This was a second goal of the study. If auditory representations are at the source of notational audiation, the most effective distractors would be events or systems that are strictly auditory in nature. On the other hand, if notational audiation involves kinesthetic-like phonatory processes, then phonatory interference conditions should cause more significant impairment.

Subject inclusion. A leading ear-training teacher (initially blind to the goals of the study) made referrals to individuals who not only demonstrated high-level abilities in general music areas (such as performance, literature, history, or pedagogy) but also demonstrated specific aural skills (related to theory and music analysis or sight-singing and dictation). These include professional orchestra players, music academy final-year students (majoring in performance, theory, or composition), advanced music undergraduates (majoring in music education or musicology), and top-honor seniors from a music-specialty high school. All prospective participants were subsequently contacted by telephone and scheduled for individual meetings. The *a priori* adopted criterion for subject inclusion was significant performance success during normal nondistracted score reading at $p < .05$ (according to a sign test). This was demonstrated by correct recognition and matching (or rejection) of at least 9 out of 12 heard tunes as original well-known embedded themes found in silently read scores. The design and test presentation used in the study mandated the participation of 18 musicians in both Experiment 1 and Experiment 2. Therefore, a large number of musicians were originally recruited and tested until a sample meeting this inclusion criterion was found. The data of the musicians who passed the critical thresholds are reported for Experiments 1 and 2. Musicians not meeting the inclusion criterion were reassigned to a control task (Experiment 3).

EXPERIMENT 1

In Experiment 1, an attempt was made to demonstrate notational audiation through the recognition of a theme embedded in an embellished phrase (read silently). The study was designed to test two predictions: (1) Task performance will be impaired under rhythmic distraction conditions (involving concurrent finger tapping with simultaneous external auditory input of an irrelevant rhythmic pattern) and (2) there will be greater performance impairment under phonatory interference conditions (involving wordless singing or humming a folk song aloud).

Method

Participants. Eighteen expert musicians were recruited from professional orchestras, music academies, university music departments, and a music-specialty high school. Six participants had completed their tertiary-level education, receiving either an Artist's Certificate or an undergraduate music degree. The participants (10 males

and 8 females) were 22 years old on average ($SD = 10.64$, range = 17–55). An average of 11 years ($SD = 4.25$, range = 5–18) of instrument tuition was reported. The most frequently reported principal instrument was the piano (55%); symphonic instruments and voice were also mentioned. All of the participants received either course credit or cash payment. Initially, 42 participants were tested, but 23 did not pass the inclusion criterion and were reassigned to Experiment 3. In addition, 1 participant was dropped from the final data set because of an electrical power failure.

Stimuli. Fifty well-known themes were selected at random by a composer (initially blind to the goals of the study) from popular symphonic or operatic sources (Barlow & Morgenstern, 1975) or contemporary Israeli folk songs (Klausner & Zur, 1988). Each original well-known theme was then embedded into a “new” embellished phrase (embedded melody) using compositional techniques such as quasi-contrapuntal treatment, displacement of registers, melodic ornamentation, and rhythmic augmentation/diminution (see Figure 1). Both the original well-known themes and the embedded melodies were recorded with a four-octave touch-sensitive MIDI keyboard controller and MIDI Sequencer (Power Traks Pro). These recordings were edited for audio glitches and converted to notation with a music composing and publishing package (Encore). Each embedded melody was recorded and transcribed in the same mode (tonality), key signature, meter, and tempo as the corresponding original well-known theme. All scores were presented to the participants as a G-clef single-line melody, with all stems pointing upward, placed in standardized measure widths, with an average melody length of 12 bars ($SD = 4.21$; range = 8–20 bars). No other markings or graphic features were present in the scores (see Figure 2).

The final scores of both original well-known themes and embedded melodies were evaluated by a music theory teacher (initially blind to the goals of the study). The evaluation was made with respect to the level of difficulty in recognizing the original well-known embedded theme. Four pairs were judged unsuitable and were dropped from the pool. The 46 remaining pairs were matched to a *melodic lure*, which was similar to the original well-known theme. These 46 melodic lures were selected from the same sources (Barlow & Morgenstern, 1975; Klausner & Zur, 1988) used for the 50 original themes. The criteria used in choosing the melodic lures were thematic or visual similarities in contour, texture, opening interval, rhythmic pattern, phrasing, meter, tonality, key signature, musical style, and (sometimes) extramusical association such as composer or musical genre (film music, march, etc.). The average length of the matched melodic lures was 12 bars ($SD = 3.76$; range = 8–16 bars). All matched melodic lures were recorded and edited in the same way as the original well-known themes and embedded melodies. The matched melodic lure for the embedded melody in Figure 1 is illustrated in Figure 3.

Forty-two triads (original well-known themes, embedded melodies, and matched melodic lures) were chosen at random from the remaining 46-triad pool. Each embedded melody was paired at random with either the original well-known theme or its matched melodic lure. This produced one set of 18 embedded melodies paired with the original well-known themes and 18 embedded melodies paired with matched melodic lures. In a second set, each embedded melody that was previously paired with a theme was paired with a lure and vice versa. Half of the participants received the first set and half received the second set. The remaining 6 triads were designated for practice trials (2 per condition).

Apparatus. Stimulus presentation and data collection was PC controlled, with a 17-in. monitor, a 16-bit sound card (Creative Labs), and two speakers. All experiments were designed and executed with MEDS 97-16 Music Experiment Development System (Kendall, 1997; Kendall & Carterette, 1992).

Design and test presentation. The experiment conformed to a single-factor within-subjects design. The experimental task re-



Figure 3. Melodic lure for the embedded theme in Figure 2 from *Suite in D* by J. S. Bach.

quired the participants to recognize themes embedded in the musical notation (read silently) and then to correctly match (or reject) a tune heard after the notation was removed from their view. Score reading took place under three conditions: (1) normal nondistracted score reading (NR); (2) rhythmic distraction (RD) conditions, consisting of simultaneously finger tapping a steady beat while hearing an additional task-irrelevant rhythmic pattern; and (3) phonatory interference (PI) conditions, consisting of wordless singing or humming a traditional Israeli folk song aloud. In total, there were 36 trials, each consisting of an embedded melody paired with either the original well-known theme or the matched melodic lure. Each block of 12 pairs represented one of the three score-reading conditions. The number of selections in the minor mode and those beginning with a *pick-up* (i.e., a partial first measure) were allocated equally between the conditions. Both condition and pair were rotated to balance biases linked to presentation order. Hence, every embedded melody was paired with both the original well-known theme and the matched melodic lure and appeared in all three score-reading conditions. In addition, the order of conditions was rotated (i.e., NR–RD–PI, PI–NR–RD, and RD–PI–NR). As a result of this procedure, 18 participants were required to complete the full rotation cycle of the experiment.

Procedure. The experiment ran for approximately 60 min and consisted of five segments, including (1) a short briefing (oral and written instructions), (2) completion of a one-page questionnaire (demographic information), (3) practice trials (2 per condition),

(4) 36 trials of the embedded melody task under the three reading conditions (12 items per condition), and (5) debriefing and remuneration. On a typical trial, each participant was exposed to the following sequence of events: Seated alongside the experimenter (who followed a written protocol), the participant was instructed to read through the scores in their entirety and not to focus on the first few measures. The words *imagine* and *imagery* were never mentioned. Then the notation of the first embedded melody appeared on the computer monitor screen and stayed in view for as long as necessary (but never for more than 60 sec). After the participant verbally acknowledged completion of the reading, the musical notation was removed from sight. A tune was immediately heard, and the participant was required to indicate as quickly as possible whether or not this heard tune was the original well-known embedded theme. The participant indicated his or her answers by pressing the “O” key for *original theme* and the “N” key for *not original theme*. A prompt appeared in the center of the monitor screen from the closure of the musical notation (i.e., onset of the heard tune) until the keypress as a reminder of the two possible response categories. In both the oral and the written instructions, the participant was told to press the appropriate key as soon as possible and not to make errors. Response times (RTs) were measured, in milliseconds, from the onset of the heard tune to the keypress. The procedures of the two interference conditions were similar to those of the normal nondistracted reading condition, with the exception of the distraction activities. During the RD condition, the participant knee-patched a steady beat



Figure 4. (A) Rhythmic distraction (RD) task: combined tapping (top staff) + rhythmic pattern (bottom staff). (B) Phonatory interference (PI) task: traditional Israeli folk song, “David, Melech Yisrael.”

while the experimenter tapped a task-irrelevant rhythmic pattern using a metal pen on the table top. The tempo of the steady beat fluctuated between trials, because it was self-generated and controlled by the participant. This combined rhythmic distractor is illustrated in Figure 4A. The same rhythmic pattern was produced for the block (all 12 items) regardless of the meter (2/2, 2/4, 3/4, 4/4, or 6/8) cited in the notation. During the PI condition, the participant repeatedly reproduced the melody of a traditional Israeli folk song (see Figure 4B). As long as the melody was audible to the experimenter, the participant could either sing without words (on a "la" syllable) or hum. The tonality key of the melody varied between trials, because it was self-generated and controlled by the participant. The same melody was sung (or hummed) for all 12 items of the block, regardless of the tonal mode (major, minor, or Dorian) cited in the notation.

Results

For each participant in each condition, the median RT for correct responses was computed. The RTs for each participant in each condition were entered into a repeated measures analysis of variance (ANOVA). As can be seen in Table 1,¹ the longest RTs were found in the phonatory interference condition. There was a significant difference between conditions [$F(2,34) = 5.37$, $MS_e = 9,856,176$, $p < .01$]. Planned comparisons indicated significantly longer RTs for the PI condition than for the other two conditions [$t(17) = 2.75$, $p < .01$ for NR vs. PI; $t(17) = 2.44$, $p < .025$ for RD vs. PI]. No significant RT difference was found between the NR and RD conditions. Furthermore, for each participant in each condition, the frequency of correct responses (i.e., an overall success rate comprised of correct hits and correct rejections) was computed. Table 1 illustrates that the highest level of task success was found in the normal nondistracted sight-reading condition. The number of correct responses for each participant in each condition was entered into a repeated measures ANOVA. There was a significant difference between conditions [$F(2,34) = 11.5$, $MS_e = 1.45$, $p < .0001$]. Planned comparisons indicated significantly more correct responses for the NR condition than for the other two conditions [$t(17) = 3.47$, $p < .001$ for NR vs. RD; $t(17) = 5.52$, $p < .0001$ for NR vs. PI]. No significant difference of success rate was found between the RD and PI conditions. Finally, the results show a decreasing level of hits and an increasing level of false alarms between the conditions, indicating

an escalating level of impairment in the sequenced order NR–RD–PI.

Discussion

The results of Experiment 1 demonstrate that expert musicians perform worse (i.e., are not as efficient in recognizing themes embedded into embellished phrases read silently) when they are distracted by phonatory activity. Because PI conditions caused significantly longer RTs than did RD conditions, it is possible to suggest that the auditory imagery triggered by musical notation elicits kinesthetic-like phonatory processes. However, since the PI task clearly generates auditory input via wordless singing or humming aloud, it could also be assumed that these effects might not be attributed to phonatory interference but to auditory distraction. In this connection, Waters and Underwood (1999) found that incoming auditory musical sequences interfered with the processing of musical notation; however, the findings do not differentiate between disruptions to processes of reading or structural analysis and disruptions to the auditory representation of musical notation. Therefore, to rule out the possibility that the interference found in Experiment 1 was due to auditory input, a second experiment was conducted.

EXPERIMENT 2

In Experiment 2, an attempt was made to examine the possibility that the auditory perception of the sung folk song alone created the effects reported in Experiment 1. To this end, in the present experiment the concurrent-singing and listening-only conditions during music score reading are compared.²

Method

Participants. A second sample of 18 expert musicians was recruited from the same pool as those in Experiment 1. Twelve had completed (or were about to complete) tertiary-level education, receiving either an Artist's Certificate or an undergraduate music degree. The participants (14 males and 4 females) were 21 years old on average ($SD = 4.01$, range = 17–25). An average of 10 years ($SD = 3.89$, range = 3–18) of instrument tuition was reported. The most frequently reported principal instrument was the piano (72%); symphonic instruments and voice were also mentioned. All of the participants received either course credit or cash payment. Initially,

Table 1
Response Times (RTs) and Number of Correct Responses in Experiment 1

Condition	RT (sec)			Correct Responses				
	<i>M</i>	<i>SD</i>	Range	Success Rate*			Hits [P(c)] [†]	FAs [P(c)] [†]
				P(c)	<i>M</i>	<i>SD</i>		
NR	9.05	3.94	4.16–16.66	.84	10.1	0.90	.82	.14
RD	8.58	5.08	3.59–23.30	.74	8.8	1.47	.78	.29
PI	11.76	6.34	5.43–27.76	.69	8.2	1.17	.78	.37

Note—NR = normal nondistracted reading; RD = rhythmic distraction; PI = phonatory interference; *M* = mean average; *SD* = standard deviation; P(c) = proportion correct. *Success rate = correct hits + correct rejections; maximum number per condition = 12. [†]Maximum number of correct hits/false alarms per condition = 6.

31 participants were tested, but 12 did not pass the inclusion criterion and were reassigned to Experiment 3. In addition, 1 participant was dropped because of excessive environmental noise during the experiment.

Design and Procedure. The stimuli, apparatus, design, procedure, and statistical methods of Experiment 2 were the same as those of Experiment 1, with the exception of a new listening-alone (LA) score-reading condition (which replaced the RD condition). In the LA condition, the participants were required to hear themselves singing a traditional Israeli folk song during score reading. At the beginning of the session, each participant sang out loud one wordless verse of the same folk song employed in the PI condition. These performances were recorded via a condenser microphone through the input channel of the PC sound card and were saved as .wav files using Wave Studio (Creative Technologies). The same individual recording was presented aloud over PC stereo speakers for all 12 items of the block via a loop playback facility, regardless of the tonal mode (major, minor, Dorian) cited in the notation.

Results

The median RT for correct responses for each participant in each condition was computed. Table 2 reveals that the longest RTs were found in the PI condition. RTs were entered into a repeated measures ANOVA, and significant differences between the conditions were found [$F(2,34) = 3.81$, $MS_e = 3,508,776$, $p < .05$]. Planned comparisons indicated significantly longer RTs for the PI condition than for the other two conditions [$t(17) = 2.15$, $p < .025$ for NR vs. PI; $t(17) = 2.02$, $p < .05$ for LA vs. PI]. No significant RT difference was found between the NR and LA conditions. Furthermore, the overall success rate for all of the participants was computed. As can be seen in Table 2, the highest level of task success was found in the normal nondistracted score-reading condition. The number of correct responses for each participant in each condition was entered into a repeated measures ANOVA. Although no significant difference was found between the conditions [$F(2,34) = 2.66$, $MS_e = 0.78$, $p < .085$], there is a clearly increasing level of false alarms between the conditions.

Discussion

The results of Experiment 2 demonstrate that expert musicians are significantly faster at recognizing an original well-known embedded theme when they are not simultaneously distracted by concurrent phonatory activity during score reading. In fact, there were no differences

in task performance when notation was read either with or without auditory input (i.e., hearing one's own singing). The results confirm the previous findings of Experiment 1, suggesting that engagement in phonatory activity was the source of interference with notational audiation. That is, the acoustical presentation and, presumably, perception of the folk song did not interfere with the musical imagery triggered by the reading of musical notation in Experiment 1.

EXPERIMENT 3

In Experiment 3, an attempt was made to highlight the experimental task itself by determining whether musicians who had performed below the 9-out-of-12 inclusion criterion would do better if the embedded melodies were presented acoustically (i.e., heard aloud) rather than visually (i.e., as notation). This question is important because of the high percentage of participants (56% in Experiment 1 and 36% in Experiment 2) who performed below criterion. That is, in Experiment 3, we explored the possibility that poor performance may have been a consequence of complexity (i.e., inappropriate experimental materials) or proficiency (i.e., underdeveloped notational audiation skills).³

Method

Participants. Twenty-three expert musicians were recruited from music academies, university music departments, and a music teachers college. The majority (87%) had been previously recruited for Experiments 1 and 2 but had failed to meet the inclusion criterion. Although one cannot be certain if the additional 13% would have met the inclusion criterion, their descriptive and demographic specifications were as close as possible to those of the original 87%. Fifteen participants had completed tertiary-level education, receiving either an Artist's Certificate or an undergraduate music degree. The participants (18 females and 5 males) were 30 years old on average ($SD = 9.94$, range = 25–55). An average of 9 years ($SD = 4.25$, range = 3–18) of instrument tuition was reported. The most frequently reported principal instrument was the piano (54%); symphonic instruments, voice, and guitar were also mentioned. All of the participants received either course credit or cash payment.

Design and Procedure. The stimuli, apparatus, design, procedure, and statistical methods of Experiment 3 were the same as those of Experiments 1 and 2, with the exception that only two conditions were implemented: the NR score-reading condition (from Experiments 1 and 2) followed by an additional block (12 trials) in

Table 2
Response Times (RTs) and Number of Correct Responses in Experiment 2

Condition	RT (sec)			Correct Responses				
	<i>M</i>	<i>SD</i>	Range	P(c)	Success Rate*		Hits [P(c)] [†]	FAs [P(c)] [†]
					<i>M</i>	<i>SD</i>		
NR	6.01	1.87	3.43–10.31	.84	10.1	1.08	.82	.13
LA	5.78	2.25	2.56–11.74	.81	9.7	0.91	.79	.20
PI	7.38	3.72	3.19–19.13	.79	9.4	1.10	.83	.26

Note—NR = normal nondistracted reading; LA = listening alone; PI = phonatory interference; *M* = mean average; *SD* = standard deviation; P(c) = proportion correct. *Success rate = correct hits + correct rejections; maximum number per condition = 12. [†]Maximum number of correct hits/false alarms per condition = 6.

Table 3
Response Times (RTs) and Number of Correct Responses in Experiment 3

Condition	RT (sec)			Correct Responses				
	<i>M</i>	<i>SD</i>	Range	Success Rate*				
				P(c)	<i>M</i>	<i>SD</i>	Hits [P(c)] [†]	FAs [P(c)] [†]
Heard aloud	8.79	5.08	3.73–17.85	.82	9.8	1.76	.87	.26
NR	12.55	3.94	6.16–25.85	.56	6.7	1.26	.65	.57

Note—Heard aloud = acoustic presentation; NR = normal nondistracted reading (i.e., notation); *M* = mean average; *SD* = standard deviation; P(c) = proportion correct. *Success Rate = correct hits + correct rejections; maximum number per condition = 12. [†]Maximum number of correct hits/false alarms per condition = 6.

which embedded melodies were presented acoustically (i.e., heard aloud).

Results

The median RT for correct responses was computed for each participant in each condition. The data were entered into a repeated measures ANOVA. Table 3 demonstrates that performance was better in the heard-aloud condition than in the NR condition. A significant difference between conditions [$F(1,22) = 11.21, MS_e = 14,506,202, p < .01$] was found. Furthermore, the overall success rate for each participant was calculated. As can be seen in Table 3, the highest level of task success was found in the heard-aloud condition. The number of correct responses for each participant in each condition was entered into a repeated measures ANOVA. There was a significant difference between conditions [$F(1,22) = 47.33, MS_e = 2.31, p < .0001$], as is illustrated by decreasing hit rates and increasing false alarm rates across conditions.

Discussion

The results of Experiment 3 demonstrate that trained musicians are significantly better and significantly faster at recognizing a tune as the original well-known embedded theme when the music is heard aloud. That is, musicians who had difficulty recognizing the theme embedded in the embellished phrase when they were presented as silently read musical notation were significantly more successful when embedded melodies were presented aloud. These results, in conjunction with those of the two previous experiments, show that although the embedded melody task is not an overly difficult one, many musicians with formal training still may not be expert enough to recognize the well-known embedded theme when the music is presented as graphic notation.

GENERAL DISCUSSION

In the present study, the participants were highly trained musicians. The sample was chosen on the basis of previous evidence that musical training develops auditory imagery skills that do not necessarily exist among nonmusicians (Aleman et al., 2000; Halpern, 1992). According to Aleman et al., the effects of musical training on auditory imagery may account for a far higher per-

centage of variance among musicians than do other components, including attention and memory (which involve enhanced ability to organize and manipulate musical information in working memory). Highly trained musicians not only play an instrument competently, but have been found to develop more efficient (i.e., faster and more accurate) skills for reading of musical notation (Kalakoski, 2001; Waters & Underwood, 1998). Whereas musicians can remember visually presented sequences of musical notes via musical imagery, nonmusicians are not able to transform visual notation into musical images. Therefore, it seems correct to assume that a fuller understanding of the nature and efficiency of notational audiation can be gained by studying expert musicians.

The experimental task developed for the study not only proved to be challenging for the musicians, but appears to have been successful in placing excessive demands on the aural skills involved in reading musical notation without disrupting visual legibility. The results seem to point out that when local surface visual cues are made ineffective in providing information needed for processing (i.e., specific visual information that musicians have become accustomed to experiencing on a daily basis), their performance success is greatly impaired. Therefore, it could be assumed that those who were able to reach the inclusion criterion did so because their notational audiation skills were developed at a higher level of proficiency than those of the participants who did not pass the threshold.

The study shows that phonatory interference impairs notational audiation. Responding was slower and false alarm rates were higher in the PI condition than in all the other conditions in Experiments 1 and 2. These findings concur with those of other studies (Calvo & Castillo, 1995; D'Esposito et al., 1996; Fleming, Goldberg, Gold, & Weinberger, 1995; Vallar & Baddeley, 1982) that found humming to be a most potent distractor and articulatory interference to cause greater disruption to phonological coding and rehearsal than either finger tapping or extraneous speech. Finally, the results show that the interference caused by the phonatory task most specifically obstructed music comprehension, whereas rhythmic activity and auditory input did not strain resources to the same extent. This result is also similar to those of studies showing that articulatory suppression both impairs comparative-judgment tasks more than does concurrent

tapping (Calvo & Castillo, 1995; Engelkamp, Mohr, & Logie, 1995) and disrupts recognition of melodic sequences (Logie & Edworthy, cited in Kalakoski, 2001).

Two models might serve as the basis for an explanation of the effects of reading musical notation under differing interference conditions. The first is founded on a model of the phonological system, and the second considers a tonological structure. In the former model, two distinct components have been described: the *phonological store*, in which representations are retained in a short-lived phonological code, and the *articulatory rehearsal process*, which serves to reactivate and maintain phonological representations. Calvo and Castillo (1995) found that some distraction tasks, such as unattended or irrelevant speech, interfere with phonological storage but allow active articulation, whereas other tasks, such as articulatory suppression, unequivocally prevent the use of the articulatory rehearsal process. Taken together, these findings suggest a two-part scheme: (1) When phonological coding is disrupted by any concurrent interference, there is a detriment in comprehension/performance; and (2) disruption causes an increase in active articulation as a means of compensating for the obstruction, and if articulation is not possible—such as during articulatory suppression or phonatory interference—then comprehension will be severely impaired. Considering the results of the present study, it seems most probable that the RD and LA conditions allowed an increase in active vocal rehearsal, whereas the PI condition did not. Hence, PI caused the greatest level of impairment. An alternative explanation can be teased out from the findings of Pechmann and Mohr (1992), who suggested the existence of a *tonal loop* (consisting of both tonal storage and rehearsal components). Their findings suggest that exposure to additional tones and sequences decreases task performance to the level of chance, whereas visual and nontonal auditory conditions result in a weaker level of interference. If this is used as a basis for an explanation of the results of the present study, then the effects of the RD and LA conditions were weaker than those of the PI condition because PI more efficiently disrupted the tonological processing and representation in imagery triggered by musical notation.

Smith et al. (1992) suggested that auditory imagery requires two mechanisms: the “inner ear” and the “inner voice.” They presented evidence showing that the former can be blocked by the presentation of irrelevant sounds, whereas the latter can be blocked by concurrent articulatory tasks. In the present study, wordless singing or humming of a melody caused interference effects that are not attributed to aural input alone (i.e., the LA condition in Experiment 2). On the one hand, performing musicians seem to develop aural skills that enable them to focus attention on their own individual parts in an ensemble while not attending to all other external sources. That is, they learn and practice placing their own musical performance in the foreground while concurrently filtering out (to a certain extent) the sounds of other in-

struments around them. On the other hand, the variance of the effects of these two conditions (LA and PI) is further evidence that notational audiation may be a skill linked more to inner voice than to inner ear.

The present study underscores significant differences that occur among musically trained individuals. About half of the musicians who participated in the study did not reach criterion during normal nondistracted score reading. Their data (see Table 3) reveal that their overall success rate (and false alarm rate) is about 50%. This finding illustrates that these musicians did not recognize the embedded melodic sources when they were presented as musical notation. In fact, their ability to differentiate between targets and lures was close to random chance level. Yet all the participants (including those in Experiment 3) had extensive formal training within one of the branches of the music profession. It appears that although all highly trained musicians attain instrumental proficiency, some have more efficient aural skills than others. To our knowledge, this distinction has not previously been reported in the experimental literature, but, rather, instrumental expertise has been acknowledged as the sole genuine hallmark of music ability and competence. One must further question if different musical instruments make particular demands in terms of motor planning and organization and, consequently, cause the development of specialized aural skills. Regarding the latter question, Brodsky, Henik, Rubinstein, and Zorman (1999) found that wind players were much more efficient in notational audiation skills than string players. Accordingly, such differences might occur either as a result of inherent demand characteristics of the instruments themselves or as a result of specific strategies and operations used over long-term practice and training regimens. Yet instrument-type differences were not identified as contributing to variance of notational audiation skill among the present sample. Rather, the internal mode of activity within the music profession comes to the surface as a new factor. For example, the majority of the participants (77%) who had previously failed to reach the criterion inclusion levels for Experiments 1 and 2 were music education and musicology majors. One possible explanation regarding this phenomenon is that these specialties require numerous hours of classroom learning, the development of countless extramusical skills (related to pedagogy, methodology, and research), and a large time investment in practicum and fieldwork. Perhaps this regimen leaves little time for daily performance practice, ear training, score singing, and dictation—all of which contribute to the general development and maintenance of aural skills such as notational audiation.

To summarize, in the present study it is proposed that notational audiation is the silent reading of musical notation resulting in auditory imagery involving kinesthetic-like phonatory processes. Therefore, these findings shed new light on previous claims (see, e.g., Sloboda, 1984) that there is no obvious way to demonstrate notational audiation and that those musicians who have this skill

are statistically rare. Regarding the first claim, in the present investigation a task that was specifically designed to demonstrate the skill—the embedded melodies task—was utilized. Regarding the second claim, whereas other researchers, such as Hallam (1995) and Banton (1995), report that approximately 10% of participating musicians had the ability to hear music from the printed page, in the study reported here we found that approximately 47% of the sample were able to demonstrate this skill. Thus, even if the former statistic is the most pessimistic estimate, it still does not reflect statistical rarity. The claim by musicians that they can hear or audiate the musical notation they read is, then, more than just a common anecdotal legend.

Finally, one might question the uniqueness of notational audiation skill.⁴ In earlier times, when only a minority of the educated could read, literacy was looked upon with awe, when in fact reading words is certainly within the potential of almost everyone. One might look on notational audiation in much the same way. Although the present study reveals that only some musicians have efficient skills, there is every possibility that a larger proportion of musicians can hear an inner voice while reading musical notation than has been thought in the past. Furthermore, just as some would argue that absolute pitch is within the potential of a larger proportion of the population than currently has the ability, musical imagery in general and notational audiation in particular may well be within everyone's grasp.

REFERENCES

- ALEMAN, A., NIEUWENSTEIN, M. R., BOCKER, K. B. E., & DE HAAN, H. F. (2000). Music training and mental imagery ability. *Neuropsychologia*, **38**, 1664-1668.
- BADDELEY, A., & LOGIE, R. (1992). Auditory imagery and working memory. In D. Reisberg (Ed.), *Auditory imagery* (pp. 179-197). Hillsdale, NJ: Erlbaum.
- BANTON, L. (1995). The role of visual and auditory feedback in the sight reading of music. *Psychology of Music*, **23**, 3-16.
- BARLOW, H., & MORGENSTERN, S. (1975). *A dictionary of musical themes: The music of more than 10,000 themes* (rev. ed.). New York: Crown.
- BEISTEINER, R., ALTENMULLER, E., LANG, W., LINDINGER, G., & DECKE, L. (1994). Musicians processing music: Measurement of brain potentials with EEG. *European Journal of Psychology*, **6**, 311-327.
- BRODSKY, W., HENIK, A., RUBINSTEIN, B., & ZORMAN, M. (1999). Inner hearing among symphony orchestra musicians: Intersectional differences of string-players versus wind-players. In Suk Won Yi (Ed.), *Music, mind, and science* (pp. 374-396). Seoul: Seoul National University, Western Music Research Institute.
- CALVO, M. G., & CASTILLO, M. D. (1995). Phonological coding in reading comprehension: The importance of individual differences. *European Journal of Cognitive Psychology*, **7**, 365-382.
- DELIÈGE, I. (1996). Cue abstraction as a component of categorization process in music listening. *Psychology of Music*, **24**, 131-156.
- DELIÈGE, I., & MELEN, M. (1997). Cue abstraction in the representation of musical form. In I. Deliège & J. A. Sloboda (Eds.), *Perception and cognition of music* (pp. 387-412). Hove, U.K.: Psychology Press.
- D'ESPOSITO, M., ONISHI, K., THOMPSON, H., ROBINSON, K., ARMSTRONG, C., & GROSSMAN, M. (1996). Working memory impairments in multiple sclerosis: Evidence of a dual-task paradigm. *Neuropsychology*, **10**, 51-56.
- ENGELKAMP, J., MOHR, G., & LOGIE, R. H. (1995). Memory for size relations and selective interference. *European Journal of Cognitive Psychology*, **7**, 239-260.
- FLEMING, K., GOLDBERG, T. E., GOLD, J. M., & WEINBERGER, D. R. (1995). Verbal working memory dysfunction in schizophrenia: Use of a Brown-Peterson paradigm. *Psychiatry Research*, **56**, 155-161.
- GODOY, R. I., & JORGENSEN, H. (Eds.) (2001). *Musical imagery*. Lisse, The Netherlands: Swets & Zeitlinger.
- GORDON, E. E. (1975). *Learning theory, patterns, and music*. Buffalo, NY: Tometic Associates.
- GORDON, E. E. (1993). *Learning sequences in music: Skill, content, and pattern. A music learning theory* (4th ed.). Chicago: GIA.
- HALLAM, S. (1995). Professional musicians' approaches to the learning and interpretation of music. *Psychology of Music*, **23**, 111-128.
- HALPERN, A. R. (1988). Mental scanning in auditory images for songs. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 434-443.
- HALPERN, A. R. (1989). Memory for the absolute pitch in familiar songs. *Memory & Cognition*, **17**, 572-581.
- HALPERN, A. R. (1992). Musical aspects of auditory imagery. In D. Reisberg (Ed.), *Auditory imagery* (pp. 1-28). Hillsdale, NJ: Erlbaum.
- HALPERN, A. R. (2001). Cerebral substrates of musical imagery. In R. J. Zatorre & I. Peretz (Eds.), *The biological foundations of music* (Annals of the New York Academy of Sciences, Vol. 930, pp. 179-192). New York: New York Academy of Sciences.
- HALPERN, A. R., & ZATORRE, R. J. (1996). Working memory in the processing of musical imagery. In B. Pennycook & E. Costa-Giomi (Eds.), *Proceedings from the Fourth International Conference on Music Perception and Cognition* (pp. 441-444). Montreal: McGill University, Faculty of Music.
- HALPERN, A. R., & ZATORRE, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, **9**, 697-704.
- HUBBARD, T. L., & STOECKIG, K. (1988). Musical images: Generation of tones and chords. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 656-667.
- HUBBARD, T. L., & STOECKIG, K. (1992). The representation of pitch in musical imagery. In D. Reisberg (Ed.), *Auditory imagery* (pp. 199-236). Hillsdale, NJ: Erlbaum.
- INTONS-PETERSON, M. J. (1992). Components of auditory imagery. In D. Reisberg (Ed.), *Auditory imagery* (pp. 45-72). Hillsdale, NJ: Erlbaum.
- JANATA, P. (2001a). Brain electrical activity evoked by mental formation of auditory expectations and images. *Brain Topography*, **13**, 169-193.
- JANATA, P. (2001b). Neurophysiological measurements underlying auditory image formation in music. In R. I. Godoy & H. Jorgensen (Eds.), *Musical imagery* (pp. 27-41). Lisse, The Netherlands: Swets & Zeitlinger.
- KALAKOSKI, V. (2001). Musical imagery and working memory. In R. I. Godoy & H. Jorgensen (Eds.), *Musical imagery* (pp. 43-56). Lisse, The Netherlands: Swets & Zeitlinger.
- KENDALL, R. A. (1997). *Music Experiment Development System (MEDS)* [Software, Version 96-32]. Los Angeles: University of California, Department of Ethnomusicology, Music Perception and Acoustics Lab.
- KENDALL, R. A., & CARTERETTE, E. C. (1992). Convergent methods in psychomusical research based on integrated, interactive computer control. *Behavior Research Methods, Instruments, & Computers*, **24**, 116-131.
- KLAUSNER, H., & ZUR, S. (1988). *Sing along with Effi Netzer: The songs we love* (Book II). Tel Aviv: Mifali Tarbut VeChinuch.
- LANGHEIM, F. J. P., CALLICOTT, J. H., MATTAY, V. S., DUYN, J. H., & WEINBERGER, D. R. (2002). Cortical systems associated with covert music rehearsal. *NeuroImage*, **16**, 901-908.
- MACKAY, D. G. (1992). Constraints on theories of inner speech. In D. Reisberg (Ed.), *Auditory imagery* (pp. 121-150). Hillsdale, NJ: Erlbaum.
- MELEN, M., & DELIÈGE, I. (1995). Extraction of cues or underlying harmonic structure: Which guides recognition of familiar melodies? *European Journal of Cognitive Psychology*, **7**, 81-106.
- MIKUMO, M. (1994). Motor encoding strategy for pitches and melodies. *Music Perception*, **12**, 175-197.
- NAKADA, T., FUJII, Y., SUZUKI, K., & KWEE, I. L. (1998). "Musical

- brain" revealed by high-field (3Tesla) functional MRI. *NeuroReport*, **9**, 3853-3856.
- PECHMANN, T., & MOHR, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory & Cognition*, **20**, 314-320.
- PETSCHKE, H. J., VON STEIN, A., & FILZ, O. (1996). EEG aspects of mentally playing an instrument. *Cognitive Brain Research*, **3**, 115-123.
- REISBERG, D. (Ed.) (1992). *Auditory imagery*. Hillsdale, NJ: Erlbaum.
- SCHNEIDER, A., & GODOY, R. I. (2001). Perspectives and challenges of musical imagery. In R. I. Godoy & H. Jorgensen (Eds.), *Musical imagery* (pp. 5-26). Lisse, The Netherlands: Swets & Zeitlinger.
- SCHUMANN, R. (1967). Musikalische Haus- und Lebensregeln [Musical house rules and maxims]. In W. Reich (Ed.), *Im eigenen Wort* [In their own words]. Zurich: Manesse Verlag. (Original work published 1848)
- SCHURMANN, M., RAIJ, T., FUJIKI, N., & HARI, R. (2002). Mind's ear in a musician: Where and when in the brain. *NeuroImage*, **16**, 434-440.
- SEASHORE, C. E. (1938). *The psychology of music*. New York: McGraw Hill.
- SLOBODA, J. A. (1974). The eye-hand span: An approach to the study of sight reading. *Psychology of Music*, **2**, 4-10.
- SLOBODA, J. A. (1984). Experimental studies in music reading: A review. *Music Perception*, **2**, 222-236.
- SLOBODA, J. A. (1994). Book review of *Language, music, and mind* by Diana Raffman 1993. *Mind & Language*, **9**, 377-385.
- SMITH, J., REISBERG, D., & WILSON, E. (1992). Subvocalization and auditory imagery: Interactions between inner ear and inner voice. In D. Reisberg (Ed.), *Auditory imagery* (pp. 95-120). Hillsdale, NJ: Erlbaum.
- VALLAR, G., & BADDELEY, A. D. (1982). Short-term forgetting and the articulatory loop. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **34A**, 53-60.
- WATERS, A. J., TOWNSEND, E., & UNDERWOOD, G. (1998). Expertise in musical sight reading: A study of pianists. *British Journal of Psychology*, **89**, 123-150.
- WATERS, A. J., & UNDERWOOD, G. (1998). Eye movements in a simple music reading task: A study of expert and novice musicians. *Psychology of Music*, **26**, 46-60.
- WATERS, A. J., & UNDERWOOD, G. (1999). Processing pitch and temporal structures in music reading: Independent or interactive processing mechanisms? *European Journal of Cognitive Psychology*, **11**, 531-553.
- ZATORRE, R. J. (1994). Musical processing in the musician's brain: Evidence for specialized neural networks. In I. Deliège (Ed.), *Proceedings of the Third International Conference on Music Perception and Cognition* (pp. 39-40). Liège: University of Liège, URPM.
- ZATORRE, R. J., & HALPERN, A. R. (1993). Effect of unilateral temporal-lobe excision in perception and imagery of songs. *Neuropsychologia*, **31**, 221-232.
- ZATORRE, R. J., HALPERN, A. R., PERRY, D. W., MEYER, E., & EVANS, A. C. (1996). Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience*, **8**, 29-46.

NOTES

1. It should be pointed out that although all the data in the study were collected and analyzed in milliseconds, the tables present the findings in seconds, for easier reading.
2. Experiment 2 was suggested by Bruno Repp of Haskins Laboratories.
3. Experiment 3 was suggested by John A. Sloboda of Keele University, England.
4. This final thought was raised by an anonymous reviewer.

(Manuscript received September 10, 2001;
revision accepted for publication December 5, 2002.)