Can Auditory Memory for Tone Pitch Be Rehearsed?

Timothy A. Keller, Nelson Cowan, and J. Scott Saults
University of Missouri—Columbia

This study investigated whether nonverbal auditory memory representations can be affected by rehearsal strategies. The comparison of the pitches of 2 tones separated by a silent, variable delay interval was examined in 2 experiments, both when participants were instructed to rehearse the pitch of the first tone covertly during the intertone interval and when such rehearsal was prevented by 1 of 2 attention-demanding distractor tasks. In both experiments, delayed tone comparison performance was superior when participants were permitted to rehearse, and the type of distractor task (verbal vs. auditory) had no effect on performance under distraction instructions. The results suggest that auditory imagery can be used strategically to slow the rate of decay of auditory information for tone pitch.

Auditory images and auditory perceptions share, at least to some extent, the same processing mechanisms (see Intons-Peterson, 1992, for a review). For instance, Farah and Smith (1983) have shown that imagining a tone of a given frequency facilitates detection of a tone at the same frequency and inhibits detection of tones at other frequencies. Crowder (1989) reported that pitch comparisons between a pure tone and a musical note were facilitated when participants were instructed to imagine the pitch of the pure tone as if it were played in the same timbre as the second, musical note. Such demonstrations of interaction between auditory imagery and auditory perception suggest that auditory imagery might provide a basis for the rehearsal of information about the modality-specific acoustic properties of sounds. Indeed, auditory imagery often is defined as a form of auditory memory (e.g., Crowder & Pitt, 1992; Intons-Peterson, 1992).

The present study is concerned with the effect of rehearsal on auditory sensory memory for tone pitch. Auditory sensory memory can be indexed by a procedure in which participants must compare two slightly different sounds presented in succession with a variable, silent interstimulus interval (Cowan, 1984). Under this delayed two-stimulus comparison procedure, a modality-specific auditory memory representation of the first stimulus (the standard) presumably must be available at the time of presentation of the second stimulus (the comparison) for participants to perform the tone comparison task correctly. The results of a large number of studies using the delayed two-stimulus pitch comparison procedure (e.g., Bachem, 1954; Bull & Cuddy, 1972; Harris, 1952; Koester, 1945; Wickelgren, 1966, 1969; Wolfe, 1886) have indicated that performance on the comparison task decreases as a function of the delay interval. Thus, the task appears to provide an index of a decaying auditory sensory memory trace. Cowan (1984, 1988) has referred to this form of acoustic storage as "long auditory sensory storage," and Massaro (1975) has called it "synthesized auditory memory."

Because participants are typically not prevented from rehearsing the standard stimulus during the intertone interval (ITI) of the delayed pitch comparison procedure, the decrease in performance as a function of ITI suggests that the loss of auditory information with time cannot be fully prevented by rehearsal. It is possible, however, that strategic rehearsal might slow but not stop the loss of auditory information. Cowan (1988) suggested that such a rehearsal process might involve the use of mental imagery to voluntarily maintain the activation of modality-specific memory elements. Similarly, Näätänen (1990, 1992) has argued, on the basis of electrophysiological evidence, that participants can voluntarily maintain an "attentional trace" for auditory sensory features by rehearsing or imagining the features relevant to a given task (Näätänen, 1992, p. 392).

Note that in both of the above descriptions of auditory sensory rehearsal, the rehearsal process is seen as involving imagination of the sensory or perceptual features of stimulation, rather than the imagination of the participant's own production of the stimuli (e.g., imagining the sound of a pure tone rather than covertly humming the pitch). This is an important distinction, because limitations of the human vocal apparatus may cause a production-based form of rehearsal to interfere with an auditory sensory memory trace in a task where very fine-grained discriminations are required. For example, Massaro (1970, Experiment 2 replication) found that the performance of participants instructed to hum the standard tone covertly during the ITI of a delayed pitch comparison task was reliably lower than that of participants instructed not to rehearse. In contrast, Wickelgren (1969), using a nearly identical procedure, reported that rehearsal instructions facili-
lated performance on a delayed pitch comparison task relative to a condition in which participants were instructed not to rehearse (although no statistical comparison was presented). One possible explanation of the discrepancy is that instructions to rehearse rather than to hum the standard tone encouraged the use of a more efficient rehearsal strategy that was based on an auditory perceptual image rather than on the covert vocal production of pitch.

It is, of course, unclear what participants do during the ITI of a delayed comparison task when they are instructed not to rehearse. A preferable manner to block rehearsal is to require that the participant perform some attention-demanding distractor task during the ITI. For instance, Deutsch (1970) presented to participants a series of spoken digits during the 5-s ITI of a delayed tone comparison task, with participants either required to attend to the digits for a serial recall task or allowed to ignore the digits. She found no significant effect of the direction of the participants’ attention. Pechmann and Mohr (1992) have used Deutsch’s (1970) procedure with individual differences in musical training as an additional factor. Pechman and Mohr found that although the manipulation of attention did not affect the pitch comparison performance of trained musicians, who performed at 96% correct or above in all conditions, the performance of nonmusicians was significantly disrupted by the requirement to attend to either spoken words or spatial stimuli during the ITI, relative to conditions in which they could ignore these interpolated stimuli.

Both Deutsch (1970) and Pechmann and Mohr (1992) also found that the presentation of a series of additional tones of different frequencies during the ITI of a delayed tone pitch comparison task produced a greater disruption of performance than did the presentation of spoken words, even when neither type of stimuli was attended. Thus, memory for pitch is undeniably subject to a stimulus-specific form of retroactive interference (see also, for example, Deutsch & Feroc, 1975; Massaro, 1970; Semal & Demany, 1991).

Studies examining the possibility of rehearsal in auditory sensory memory have included stimuli in the ITI, so they cannot reveal the effects of rehearsal in the absence of interference. In contrast, in the present study the interference during the ITI consisted only of the imagination of previously presented stimuli (tone or word sequences). An additional question we addressed is whether the specificity of interference observed in previous studies extends to a situation in which the interpolated stimuli are not actually presented during the ITI, but are only imagined. For example, a task in which a participant recalls a series of tones of different frequencies presumably requires the participant to imagine the pitches of the tones. If auditory images are functionally isomorphic with auditory perceptions, then the performance of this task during the ITI of a delayed comparison task would be expected to produce a greater disruption of tone comparison performance than a task requiring the recall of a series of digits.

It should be noted that, with the exception of Massaro (1970), all of the studies discussed above have compared performance under rehearsal and nonrehearsal conditions at only one delay interval. As a result, they assessed only the strength of memory for pitch after a fixed interval of time had elapsed since the presentation of the standard tone and could not be used to evaluate the effect of rehearsal on the rate of information loss. Unequal memory loss rates could be inferred if it can be assumed that participants perceived the tones equally well in both the rehearsal and nonrehearsal conditions. However, all of these studies used blocked presentation of rehearsal and nonrehearsal trials, which permitted that participants might not have attended to the tones to be compared equally well across conditions. Therefore, any effect of the rehearsal condition on performance in these studies could be attributed to differences in perceptual processing of the tones rather than to differential rates of tone pitch memory loss.

Our present experiments reexamine the question of whether some form of strategic rehearsal can be used to maintain auditory sensory information about tone pitch by using a task in which participants compare the pitches of two slightly different tones separated by a variable, silent ITI. The role of perceptual processing was addressed in Experiment 1 by using two delays (0.5 and 10.0 s) and in Experiment 2 by intermixing the rehearsal and interference trials in common trial blocks.

To avoid ceiling and floor effects in the data, the difference in frequency between the tones to be compared was individually adjusted for each participant to a level that allowed 71% correct performance at a 5-s ITI. The 71% level was not chosen arbitrarily but has been shown theoretically to be the result of an efficient adaptive tracking procedure in which the discrimination is made easier if the participant responds correctly on two successive trials and is made harder if the participant responds incorrectly on a trial (Levitt, 1971).

In both of the experiments, we compared performance for a condition in which the participants were permitted to rehearse the first tone silently, in any manner they wished, with conditions in which rehearsal was prevented by a silent distractor task involving the immediate recall of auditory or verbal information during the retention interval. In each of the three conditions, the participants simultaneously received a series of tones and visually presented digits prior to the presentation of the tones for the comparison judgment. Thus, the differences among the conditions involved only the participants’ attentional set (attend to auditory distractors, attend to verbal distractors, or ignore both types of distractors) and the corresponding intertone task that the participant was required to carry out during the delay interval between the standard and comparison tones (recall auditory distractors, recall verbal distractors, or rehearse the standard tone, respectively).

The difficulty of the distractor tasks was equated across conditions by determining each participant’s memory span for each type of information, so that any differences between conditions in tone comparison performance could be interpreted as resulting from the type of information processed during the retention interval rather than from the difficulty of the distractor task. If differences among the three conditions affect only the perceptual processing of the tones, rather than the rate of tone information loss, then only a main effect of condition is expected. On the other hand, if the rate of pitch information loss is affected by the manipulation of the type of processing carried out during the ITI, then a Condition × ITI interaction is expected as well.
Experiment 1

Method

Participants

The participants were 24 undergraduate students (12 men and 12 women) who had no known hearing problems and were free of colds at the time of testing. None of the participants were professional musicians and none claimed to possess absolute pitch. They were recruited from introductory psychology courses and received course credit in return for their participation.

Design

We first tested the participants to determine their 71% correct Δf threshold for the tone comparison task at an intermediate, 5.0-s ITI, and this Δf level was used in all further phases of the experiment. Each participant was tested in three blocked intertone task conditions: verbal distraction, auditory distraction, and rehearsal. The order in which these conditions were tested was completely counterbalanced across each subset of 6 participants. Within each condition the participant received 32 test trials. On half of these trials the ITI was 0.5 s and on the remaining trials the ITI was 10.0 s. Within each test block, the 16 trials at each ITI were randomly intermixed.

Apparatus and Stimuli

Participants were tested one at a time in a sound-attenuated chamber. The experiment was controlled by an Apple Macintosh II microcomputer located outside the chamber. The auditory stimuli were created with a sound editing program and were digitally stored on the hard disk of the microcomputer. The auditory stimuli consisted of seven different sets of pure tones (sinusoidal waves). Each of these seven sets contained nine different tones that were created so that the difference in frequency between the standard and comparison tones (Δf) could equal 1.25, 2.5, 5.0, 10.0, 20.0, 40.0, or 80.0 Hz. Each of the seven Δf sets was centered at 405.0 Hz, so, for instance, the 1.25-Hz set contained nine tones varying in frequency between 400.0 and 410.0 Hz in 1.25-Hz steps, whereas the 80.0-Hz set contained nine tones varying in frequency from 85.0 to 725.0 Hz in 80.0-Hz steps. Each tone was 200 ms in duration, with 20-ms rise and fall times. The standard and comparison tones on a given trial always consisted of adjacent tones from one of the seven sets, resulting in eight possible tone pairs that could be presented with either the higher tone or lower tone first. The auditory stimuli were output through the computer's digital-to-analog converter and were presented through audiological headphones at 80 dB(A).

The verbal distractor stimuli were visually presented on a 19-inch AppleColor high-resolution RGB monitor as black images on a white background. For the verbal distraction condition, the to-be-remembered stimuli were drawn from the digits 1–9. They were approximately 5 mm high on the screen and were presented sequentially in boxes at discrete locations moving from left to right across the top of the screen. Each digit was visible for 500 ms, with the offset of a digit followed immediately by the appearance of the next digit in the list.

For the auditory distraction task, the to-be-remembered stimuli were drawn from the same set of nine tones used for the tone comparison task. The onset of each tone in a sequence was synchronized with that of the corresponding digit, so that stimulus onset asynchrony for both types of stimuli was 500 ms, resulting in a 300-ms interstimulus interval between successive tones in a list and a 0-ms interstimulus interval between successive verbal stimuli.

A total of 32 lists each was created for each of the two distractor task conditions. The stimuli at each serial position in a list was randomly selected from the set of nine possible values with the additional constraints that no two successive stimuli in a list were allowed to be equal to each other, or adjacent to each other within the set of nine possible values. In addition, the auditory lists were constructed such that the tones to be used as the standard and comparison tones on a trial were not included in the list of distractors for a given trial. Because the lists were created independently for each of the conditions, there was no systematic correspondence between the digit and the tone being presented at a given time. The 32 lists for a condition were presented in a different random order to each participant.

The participants entered responses by manipulating a mouse that controlled the position of the cursor on the screen, and by depressing a button on the mouse when the cursor was at the location where they wanted their response to be recorded. Responses were recorded by the program as the number of pixels from the top of the screen to the location of the cursor at the time the participant depressed the mouse button.

Procedure

Each participant completed a single session that lasted approximately 2 hr. The participants first received practice at performing the tone comparison task alone. A total of 32 lists of 10 items each was created for each of the two distractor tasks in isolation: practice on one of these conditions was immediately followed by a span test for the condition. Finally, the participants received 16 trials of practice and 32 test trials for each of the three instructional conditions. Testing for a given condition immediately followed the practice trials for that condition.

Tone comparison practice and Δf pretesting. At the beginning of the session, the participants first completed 16 trials of practice on the tone comparison task alone. Of these trials, 8 had a 0.5-s ITI and the remaining 8 had a 10-s ITI, with the two ITIs randomly intermixed. On each trial, the participants depressed the mouse button to hear the standard tone, followed by a blank ITI of either 0.5 or 10.0 s, followed by the comparison tone. The comparison tone was always 10 Hz higher or lower than the standard tone during this initial practice phase. Following offset of the comparison tone, two response boxes appeared side by side on the screen, labeled higher and lower, and the participant used the mouse to select the appropriate response.

The initial practice phase was followed by the Δf test procedure during which the ITI was always 5 s. To estimate the participant's Δf threshold, the program that controlled the experiment used an adaptive tracking procedure (Levitt, 1971) that varied the difference in pitch between the standard and comparison tones. The initial Δf level was 40.0 Hz. Whenever the participant gave two consecutive correct responses, the Δf for the next trial was decreased to the next lower value (i.e., new Δf = previous Δf/2). Whenever the participant erred on one trial, the Δf for the next trial was increased to the next higher value (i.e., new Δf = previous Δf + 2). This "2-down, 1-up" adaptive tracking procedure theoretically allows estimation of the stimulus level corresponding to 71% correct performance (Levitt, 1971), and this procedure is widely used for determining difference thresholds in perceptual research. The program kept track of the reversals, or the Δf values at which a transition between increases and decreases in the Δf value occurred. The first reversal was discarded, and the geometric mean of the next six reversals was taken as the estimate of the participant's 71% correct Δf threshold at the 5-s ITI. The closest of the 7 preselected Δf sets to this mean was then used in further testing for that particular participant.

Distractor task practice and span testing. Each participant completed 16 trials of practice on each of the two distractor tasks, with the span test for a condition immediately following the practice trials for that condition. The order of practice and testing during this phase of the experiment was identical to the order in which the conditions would be tested for the participant later in the experiment.
Each trial of practice contained four to-be-remembered stimuli, along with four stimuli from the other stimulus type. We first instructed the participants in the use of the mouse and told them that they would initiate each trial by moving the mouse so that the cursor was within an area of the screen marked Begin and by depressing the mouse button.

For the verbal distraction condition, we told the participants to attend only to the digits and to try to remember the order in which they appeared. Following presentation of the four digits, a response screen was displayed that contained a vertically oriented rectangle at the horizontal location where the first stimulus was presented (see Figure 1). We told the participant to assume that the top of this rectangle represented nine and the bottom represented one. The participant’s task was to move the cursor to the vertical position within this rectangle that would represent the first digit, and to depress the mouse button to record the response. This resulted in a black bar being displayed at the position where the response was made, and in the rectangle being moved to the next position to the right so that the next response could be made. Following the participant’s final response, feedback was provided in the form of a series of white bars showing the correct positions for each stimulus in the series, superimposed over the participant’s own responses. We instructed the participants to note any discrepancies between the responses made and the correct positions.

For the auditory distraction condition, we instructed the participants to try to remember the pitches of the four tones, and to then reproduce them using the same response screen as before, with higher-pitched tones being represented closer to the top of the rectangle and lower-pitched tones being represented closer to the bottom. We told the participants that they were not expected to be able to reproduce the pattern perfectly but that they should try to capture the relative pattern of the pitches in the sequence of tones.

Immediately following completion of the 16 trials of practice for a distractor task condition, we tested the participants for their memory span for that condition. We told the participants that their task was the same as in the practice for that condition but that the number of stimuli presented would increase on successive trials. The task began with three to-be-remembered items, and another item was added after each trial in which they correctly reproduced the entire list. If the participant failed to reproduce all the items correctly on a trial, the following trial used the same number of stimuli again. Testing continued until the participant missed two trials at a given series length. The procedure was then repeated, beginning again with three items, and the longest list length accurately reproduced in either run was taken as an estimate of the participants’ memory span for that condition.

The scoring of the participant’s responses was carried out within the program that ran the experiment. The response at the first serial position was always considered correct, but each subsequent response was considered correct only if it was higher than the previous response and the actual stimulus was higher than the actual previous stimulus, or if it was lower than the previous response and the actual stimulus was lower than the actual previous stimulus. As a result, a perfect score at a given list length did not require the participant to make correct absolute judgments of each stimulus, but rather, only required the participant to reproduce the relative changes in stimulus level. Although responses were scored in the same manner for both of the distractor conditions, we did not tell the participants the details of the scoring procedure for the verbal condition, so from their point of view the task required absolute memory for each digit in the sequence. For the auditory condition, we informed the participants that their responses would be scored correct as long as they captured the correct contour of the pitches.

Dual task practice and testing. After the participant’s memory span for each of the distractor task conditions had been determined, each participant completed 16 trials of practice and 32 test trials in each of the three conditions. For the practice trials, we told the participants that they would now practice combining two of the tasks. When they depressed the mouse button, a series of digits and tones was presented, with the number of items presented being determined by the participant’s score on the earlier memory span test for that condition. We told the participants to attend to the appropriate stimuli. For the rehearsal condition, we told the participants to ignore the distractor items and to concentrate only on the tone comparison task. The number of distractor items that preceded the tone comparison task in the rehearsal condition was set at the same level as for the auditory condition. Following the last item of the distractor task list, the screen went blank, and the participant heard the two stimuli for the tone comparison task. We informed the participants that if there was only 0.5 s between the standard and comparison tones, they would not have to reproduce the distractor task stimuli to which they had been told to attend, but if there was 10.0 s between the two tones, the response screen would appear and they should reproduce the distractor task series as many times as possible before the comparison tone occurred. After the comparison tone was presented, the participants were required to make the higher-lower judgment as before. We told the participants that they should consider the two tasks (tone comparison and memory list reproduction) to be of equal importance. Following the tone comparison response, the computer provided feedback on whether the participant’s higher-lower decision was correct, as well as feedback on the number of correct responses on the distractor task.

Results and Discussion

Δf Test Performance

The mean estimated Δf level required for 71% correct performance at a 5.0-s ITI was 14.06 Hz (SD = 11.20). For individual participants, the Δf level used for testing ranged from 2.5 to 40.0 Hz. The high degree of intersubject variability is typical for tone comparison data (e.g., see Green, 1976, p. 262).

Span Test Performance

The mean memory spans for the verbal and auditory distractor task conditions are shown in Table 1. The partic-
participants could recall significantly more digits than tones, as indicated by a one-sample paired t test, $t(23) = 23.387$, $p < .0005$, two-tailed. This finding is consistent with the idea that the participants were using a more efficient form of coding to maintain the digits.

**Distractor Task Performance**

The means and standard deviations for the total distractor task responses per trial and the proportion of correct responses per trial for each of the distractor task conditions are shown in Table 1. Because the participants were required to repeat the distractor task sequence as many times as possible during each 10.0-s ITI, it was possible for the total number of responses on a trial to be greater than the number of items presented to the participant. The total responses per trial therefore provides a measure of the rate at which the participants were performing the task, and the proportion of correct responses provides a measure of accuracy. There was no reliable difference between the two conditions in the speed of distractor task responding, but the participants were significantly more accurate in the verbal condition, paired $t(23) = 3.400$, $p < .005$, two-tailed.

The fact that the participants were more accurate in performing the verbal distractor task than the auditory distractor task suggests that the procedure of adjusting the number of distractor task items presented in each condition on the basis of a prior memory span test was not completely successful in equating the difficulty of the intertone tasks. It is possible that the participants benefited more from additional practice on the verbal task than from practice on the auditory task. Alternatively, it could be argued that this finding suggests that the participants were investing less attention in the intertone task during the auditory distraction block than during the verbal distraction block.

**Tone Comparison Performance**

Figure 2 shows the mean proportion of correct tone comparison responses at each ITI for the three instructional conditions. These data were submitted to a 3 (instructional condition) × 2 (ITI) repeated measures analysis of variance (ANOVA), which revealed a significant interaction between these two variables, $F(2, 46) = 4.721$, $MSE = 0.006, p < .05$, as well as a significant main effect of ITI, $F(1, 23) = 82.386$, $MSE = 0.013, p < .0001$. The main effect of instructions was not significant. The simple main effect of instructional condition was not significant for the 0.5-s ITI but was significant for the 10.0-s ITI, $F(2, 46) = 5.089$, $MSE = 0.009, p < .05$. Tukey's honestly significant difference (HSD) pairwise comparisons indicated that the mean proportion correct in the rehearsal condition at the 10.0-s ITI was significantly greater than that in either of the distraction conditions, $ps < .05$, which did not differ significantly from each other.

The results of Experiment 1 suggest that when the difficulty of the tone comparison task was individually adjusted for each participant, both verbal and auditory mental processing carried out during the ITI resulted in significantly poorer retention of the standard tone relative to a condition where the participant was free to rehearse the standard tone. The finding that the type of distractor task had no effect on performance at a 10.0-s ITI is consistent with the view that a silent distractor task has its effect by preventing strategic rehearsal rather than by directly interfering with an auditory sensory representation of the standard tone.

The results of Experiment 1, however, cannot completely rule out an interpretation in terms of an effect of distraction on perceptual processes rather than on a memory retention process. Because the instructional conditions were blocked in this experiment, it is possible that attention to the standard tones, comparison tones, or both was greater in the rehearsal condition. Although the lack of a significant effect of instructional conditions at the 0.5-s ITI in Experiment 1 suggests that the participants were perceiving tones equally well across conditions, it is possible that the tone comparison task was not sensitive enough to reveal differences among the conditions in the perceptual processing of the tones at such a brief delay interval. Experiment 2 examines this question by comparing performance between distraction and rehearsal conditions in a situation where the participants do not know whether they will have to recall the distractor items until after presentation of the standard tone.

**Table 1**

<table>
<thead>
<tr>
<th>Distractor task</th>
<th>Memory span</th>
<th>Total responses per 10-s ITI</th>
<th>Proportion of correct responses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Verbal</td>
<td>6.54</td>
<td>1.22</td>
<td>5.20</td>
</tr>
<tr>
<td>Auditory</td>
<td>4.71</td>
<td>1.37</td>
<td>4.96</td>
</tr>
</tbody>
</table>

*Note.* ITI = intertone interval.

Experiment 2

In this experiment, rehearsal and distraction conditions were randomly intermixed within a single block of test trials. However, because the participants could not be expected to attend to more than one type of distractor sequence at a time, it was necessary to block the distractor task conditions (i.e.,
attend verbal series vs. attend auditory series). The signal indicating whether the participants would have to recall the distractor items was presented only after presentation of the standard tone. This arrangement ensures that participants give equivalent attention to the distractor items and standard tone across the intertone task conditions (i.e., distraction vs. rehearsal). The difference in tone comparison performance between rehearsal and distraction found in the previous experiment should be replicated here only if it is due to an effect on retention rather than on perception.

A second procedural change between Experiments 1 and 2 involves the timing of the offset of the response screen used by the participants to record responses to the distractor tasks, and the timing of the onset of the comparison tone. In this experiment, the response screen was removed from the screen 0.5 s prior to presentation of the comparison tone in order to cue the participant to prepare for presentation of the comparison tone.

Finally, in addition to these mixed intertone task blocks, a blocked condition in which the participants knew they would always be allowed to rehearse the standard tone if the ITI was 10.0 s was included. Comparison between performance under this condition and on the rehearsal trials of the mixed-block condition allows for an assessment of the contribution of attentional set to rehearsal performance.

Method

Participants

Twenty-four undergraduate students (13 men and 11 women), who met the same requirements as the participants in Experiment 1, completed a single 2-hr session. The participants received course credit in exchange for their participation.

Design

The design of this experiment differed somewhat from that of the previous experiment (and, as a result, the analyses differed). The participants completed four blocks of test trials, each of which consisted of an equal number of 0.5-s and 10.0-s ITIs, randomly ordered within the block. The first and last test blocks completed by each participant were performed under rehearsal-only instructions identical to those for the rehearsal condition in Experiment 1. Each of these blocks consisted of 16 trials, 8 with a 0.5-s ITI and 8 with a 10.0-s ITI. For Blocks 2 and 3, the participants were instructed to attend to either the digit or tone series throughout the block, with the distractor task condition order counterbalanced across the participants. Test Blocks 2 and 3 each consisted of 48 trials, composed of 24 trials at each ITI. For the trials at the 10.0-s ITI, 16 required the participant to reproduce the contour of the attended distractor sequence during the ITI and 8 required the participant to rehearse the standard tone during the ITI.

As mentioned above, in Blocks 2 and 3 of testing, the participants were required to attend to one of two types of distractor series on each trial and were required to either rehearse the standard tone or perform the distractor task when the ITI was 10.0 s. Because the cue signalling the intertone task was not presented on 0.5-s trials within these blocks (i.e., no intertone task was specified), an ANOVA involving both ITIs was not possible. Consequently, only the 10.0-s data are included in the main analysis. If the requirement to attend to the distractor stimuli was responsible for the lower level of performance under distraction instructions in the previous experiments, then no effect of intertone task condition should be expected in the 10.0-s data.

Procedure

The initial phases of the experiment followed the procedures used in Experiment 1. The participants began the session with 16 trials of practice on the tone comparison task, followed by the Δt test. They then received 10 trials of practice on one of the two distractor tasks, followed immediately by the memory span test for that distractor task, and then they completed the practice and memory span test for the other distractor task. The order of the distractor task conditions was counterbalanced across participants and was maintained in the order of testing for each participant later in the session.

Following memory span testing for the second condition, all participants performed one block of 16 test trials according to the rehearsal instructions used in the previous experiment. As in the previous experiment, in this condition the number of auditory and visual distractor items presented prior to the standard tone was individually set at the participant's memory span for the auditory distractor task condition. We told the participants to ignore both visual and auditory distractor lists presented prior to the standard tone. In contrast to the previous experiment, during the block of rehearsal-only trials the words Rehearse the last tone were presented on the screen 0.5 s after offset of the standard tone.

After completing one block of rehearsal-only trials, the participants practiced combining one of the distractor tasks with the tone comparison task. The procedure for this practice block followed that of Experiment 1, except that on 4 of the 12 10.0-s ITI trials, the rehearsal cue (the words Rehearse the last tone) was presented on the screen 0.5 s after offset of the standard tone. We told the participants that if this message appeared, they were to forget the distractor items and instead concentrate on remembering the standard tone during the remainder of the 10.0-s ITI. On the remaining 8 trials with a 10.0-s ITI, the response screen for the distractor task appeared 0.5 s after offset of the standard tone. We told the participants that if this response screen appeared, they were to reproduce the distractor item sequence as many times as possible during the remainder of the 10.0-s ITI. On the remaining 12 trials of this practice block, the ITI was 0.5 s, and neither the rehearsal cue nor the distractor task response screen appeared during the ITI. Note that these trials with a 0.5-s ITI could not be considered purely rehearsal trials because the participants did not learn that they could forget the distractor series until the comparison tone arrived.

Dual-task testing for the participant's first distractor task condition immediately followed practice for that condition. The procedure was the same as that for the practice block, except that the test block included 48 trials consisting of 24 trials with a 0.5-s ITI, 16 trials with a 10.0-s ITI and the requirement to perform the distractor task, and 8 trials with a 10.0-s ITI and the requirement to rehearse the standard tone.

Following the test block for the first distractor task condition, we gave the participants a 24-trial block of practice at combining the other distractor task with the tone comparison task. This was followed by the test block for the participant's second distractor task condition. Finally, the participants completed a second block of 16 rehearsal-only trials. This block was conducted according to the same procedure as that for the first block of test trials. At the conclusion of the experiment, the experimenter asked all participants to describe what strategies they used to remember the digit distractor series and the tone distractor series on the dual-task test blocks, as well as to describe any strategies they used for remembering the standard tone on trials for which rehearsal was the intertone task.
Results and Discussion

Performance Measures

$\Delta f$ test performance. The mean $\Delta f$ estimate for 71% correct performance at a 10.0-s ITI was 10.45 Hz ($SD = 7.35$). For individual participants, the $\Delta f$ levels used for testing ranged from 2.5 Hz to 20 Hz.

Span test performance. The mean spans for the verbal and auditory distractor task conditions are shown in Table 2. As in the previous experiment, the participants could recall significantly more digits than tones, paired $t(23) = 3.06, p < .01$, two-tailed.

Distractor task performance. Table 2 also displays the means and standard deviations for the total distractor task responses per trial and the proportion of correct responses per trial in each of the two distractor task conditions during dual-task testing. As in the previous experiment, the difference between conditions was not reliable for the number of responses made, but the accuracy of the responses was significantly greater for the verbal distractor task, paired $t(23) = 4.096, p < .0005$, two-tailed. Again, this finding indicates that the difficulty of the distractor tasks was not perfectly equated across conditions.

Tone comparison performance. Figure 3 displays the mean proportion of correct tone comparison responses at each ITI for each combination of block type and intertone task. A preliminary test of the effect of block for the data from the rehearsal-only condition indicated no reliable difference in overall proportion correct across the two ITIs between Block 1 of testing ($M = .777, SD = .129$) and Block 4 of testing ($M = .786, SD = .179$), paired $t(23) = .394, p > .6$, two-tailed. The data from these separate blocks were therefore combined in subsequent analyses.

To determine whether an effect of attentional set on the perceptual processing of the tones to be compared could account for the superior performance at the 10.0-s ITI under rehearsal instructions in the previous experiment, an ANOVA with the variables block type (attend verbal, attend auditory, and ignore both) and ITI (0.5s and 10.0s) was carried out for the data from all trials in which the participants were not required to perform a distractor task during the ITI. Because the participants were never required to perform the verbal or auditory distractor task when the ITI was 0.5-s, all test data at this ITI were included in the analysis. At the 10.0-s ITI, in the attend-verbal and attend-auditory blocks, only data from trials in which we instructed the participants to rehearse the standard tone during the ITI were included.

This analysis of trials with no intertone distraction revealed a significant main effect of block type, $F(2, 46) = 3.30, MSE = 0.010, p < .05$. One can see from Figure 3 that the performance in rehearsal-only blocks was better than that on trials with no intertone task from other blocks. Follow-up pairwise comparisons (Tukey’s HSD) indicated that only the difference between performance across the two ITIs under the rehearsal-only block versus under the auditory distraction block was reliable, $p < .05$. As expected, the effect of ITI was highly significant, $F(1, 23) = 19.021, MSE = 0.013, p < .0005$. The effect of block type did not, however, interact with that of ITI, $F(2, 46) = 0.104, MSE = 0.012, p > .8$. Thus, it appears that the requirement to attend to series of distracting tones prior to presentation of the standard tone can disrupt tone comparison performance even when the distracting tones do not have to be recalled during the ITI. The absence of a Block Type x ITI interaction in the present analysis, however, suggests that this effect of attentional set does not influence auditory memory retention. Rather, the disruption of performance resulting from attention to a distracting series of tones presented before the standard tone appears to be confined to perceptual processing of the standard tone and does not affect the rate at which the availability of information about the pitch of the standard tone is lost.

The data displayed in Figure 3, when inspected across blocks for only the 10.0-s ITI, also suggest that the effect of attentional set on perceptual processing of the tones to be compared cannot completely account for the difference between rehearsal and distraction instructions found in the previous experiment. It appears that a substantial difference in performance was still obtained in a comparison between rehearsal trials and distraction trials even when the participants were required to attend to the distracting stimuli before the onset of the standard tone under both rehearsal and distraction conditions (see Figure 3). This suggestion was examined statistically in several ways.

First, the proportion-correct data for only the 10.0-s ITI from Experiment 2 were submitted to a one-way repeated measures ANOVA with five levels of the condition variable representing the five possible combinations of block type and intertone task (attend auditory and recall auditory, attend verbal and recall verbal, attend auditory and rehearse standard, attend verbal and rehearse standard, and ignore both and rehearse standard). This analysis revealed a significant main effect of condition, $F(4, 94) = 9.869, MSE = .011, p < .0001$. As in Experiment 1, post hoc pairwise comparisons (Tukey’s HSD) indicated that both the verbal and auditory distractor tasks resulted in reliably lower performance than the blocked rehearsal-only condition, $p < .01$, but the two distraction conditions did not significantly differ from one another. In addition, the means under verbal distraction and under auditory distraction were both reliably lower than those for the rehearsal trials carried out within the same test blocks (the attend auditory and rehearse standard, and the attend verbal and rehearse standard conditions), $p < .01$. This latter result clearly indicates that the requirement to perform a silent distractor task during the 10.0-s ITI did affect retention of the standard tone.

In addition, to perform an analysis comparable to that conducted for the previous experiment, data from both ITIs
were submitted to a separate $3 \times 2$, Block Type (rehearsal-only, auditory distraction, verbal distraction) $\times$ ITI (0.5-s, 10.0-s) repeated measures ANOVA. This analysis excluded the data from rehearsal trials under the two distraction attentional sets because the 0.5-s data could not be separated into rehearsal versus distraction trials. Consistent with the results of Experiment 1, this analysis revealed a significant main effect of ITI, $F(1, 23) = 65.546, MSE = .013, p < .0001$, a significant main effect of block type, $F(2, 46) = 26.691, MSE = .006, p < .0001$, and a significant ITI $\times$ Block Type interaction, $F(2, 46) = 7.72, MSE = .006, p < .005$ (see Figure 3).

In contrast to the results of Experiment 1, a test of the simple main effect of block type at the 0.5-s ITI did indicate a reliable difference among the means in this experiment, $F(2, 46) = 4.08, MSE = .006, p < .05$ (see Figure 3). Post hoc Tukey's HSD pairwise comparisons indicated that only the difference between the rehearsal condition and the auditory distraction condition was significant, $p < .05$. It is unclear why an effect of attentional set on tone comparison performance at the 0.5-s ITI was present in this experiment when no such effect was found in Experiment 1. Perhaps not knowing if a series of tones will have to be recalled (as in this experiment) is somewhat more of a distraction from tone perception than is knowing for certain that it will have to be recalled (as in the previous experiment). Nevertheless, the pairwise comparisons reported above for the 10.0-s ITI data indicate that performance on the blocked rehearsal condition was reliably greater than that on the two distractor task conditions, which did not significantly differ from each other. The ITI $\times$ Block Type interaction must therefore be interpreted as resulting from a greater effect of block type at a 10.0-s ITI than at a 0.5-s ITI, and this pattern is consistent with that found in Experiment 1.

**Individual Differences in Self-Reported Strategies**

In this experiment, we systematically questioned the participants about what they were doing to remember the tones and digits during the course of the experiment. In response to the question about how they remembered the pitch of the standard tone under rehearsal instructions, all 24 of the participants reported using some form of auditory imagery. No participant reported using any type of visual label to encode the standard tone. Twenty of the participants described their strategy in purely perceptual terms (e.g., "I imagined the sound of the tone over and over in my head"), but the remaining 4 participants described the form of imagery in terms of a self-produced sound (e.g., "I hummed the note to myself").

In response to the question about how the digit sequence was retained for the verbal distractor task, all participants reported a strategy of covert verbal rehearsal (e.g., "I said the numbers over and over to myself"). For the auditory distractor task, however, clear differences in self-reported strategies did emerge. Eleven of the participants maintained that they used some form of auditory imagery to recall the pattern of the pitches (e.g., "I just heard the sounds in my head," or "I hummed the tones like a song"). The other 13 participants claimed that they recoded the tone series into another form of representation at the time of presentation in order to perform the distractor task. Seven of these participants described their recoding strategy in visuospatial terms (e.g., "I thought about where the bar should go for each note as I heard it, and then I just remembered how the bars went"). The remaining participants claimed that they used an explicitly verbal form of coding (e.g., "I just said to myself 'higher' or 'lower' as I heard each tone and I remembered the pattern that way").

To examine whether this difference in the participants' strategies for performing the auditory distractor task affected performance on our dependent measures, the participants were divided into two groups on the basis of whether they reported using auditory imagery (11 participants) or some form of recoding to perform the auditory distractor task (13 participants). Comparisons of the performance of these two groups of participants on our three measures of auditory distractor task performance (memory span, number of responses per trial, and proportion of correct responses) indicated that only the proportion of correct auditory task responses during the dual-task test block differed between the groups, $t(22) = 2.240, p < .05$, two-tailed. The mean proportion of correct auditory distractor responses for the participants who used an auditory imagery strategy was .775
REHEARSAL OF TONE PITCH

(SD = 0.91), and for those who used a verbal or spatial recoding strategy it was 0.688 (SD = 0.096).

The fact that the participants differed in the form of mental processing used for the auditory distractor task allows for another opportunity to examine whether auditory imagery of the distracting sequence directly interferes with an auditory modality-specific representation of the standard tone, or simply prevents rehearsal of the standard tone. If imagining a series of tones has an effect on memory for pitch similar to that of actually hearing a series of interfering tones, it can be predicted that maintenance of the standard tone will be hurt by the auditory distractor task more in participants processing the distractor sequence through an auditory imagery strategy than in participants using a verbal or spatial recoding strategy.

Contrary to these predictions, though the auditory distractor task processing strategy was found to affect the level of tone comparison performance in auditory distractor blocks, the effect was not specific to distractor trials. Instead, the effect of distractor strategy occurred equivalently for rehearsal and distractor trials within the auditory distractor trial blocks. Means for the 10-s ITI for the participants who used the two strategies are shown in Table 3. An ANOVA was conducted for the 10-s ITI trials within auditory distractor blocks with condition (rehearsal vs. distraction) as a within-subject variable and distractor strategy (auditory imagery versus recoding) as a between-subjects variable. This ANOVA revealed main effects of the condition, F(1, 22) = 6.58, MSE = .09, p < .02, and the strategy, F(1, 22) = 6.48, MSE = .14, p < .02, but no interaction between the two, F(1, 22) = 1.96, MSE = .03, p > .1. Table 3 shows that there was an advantage for rehearsal trials and for the participants who used a recoding strategy to perform the distractor task, but no group difference specific to the distractor trials.

The difference between the participants using different strategies in the auditory distractor blocks theoretically could reflect either a direct effect of the strategy difference or some more general difference between these groups of participants. However, an examination of these same participants’ performance in other conditions revealed no strategy-group differences (see Table 3). Neither an ANOVA of the verbal distractor blocks with the same variables as the ANOVA of the auditory distractor blocks, nor a one-way ANOVA of the rehearsal-only blocks, nor separate ANOVAs of the 0.5-s data in the three types of trial blocks showed any effects of strategy group approaching significance. Therefore, the group difference seems specific to auditory distractor trial blocks. Given that it occurs even on rehearsal trials within those blocks, it appears to reflect an effect of early distractor processing on the clarity with which the standard tone can be encoded for later comparison, rather than an effect of distractor processing strategy during the ITI on forgetting of the standard tone.

General Discussion

The primary purpose of the present study was to determine if auditory imagery could provide a basis for the rehearsal of the detailed acoustic information contained in an auditory modality-specific sensory memory representation. The results of the two experiments reported here are consistent in indicat-

<table>
<thead>
<tr>
<th>Block type</th>
<th>Auditory imagery(a)</th>
<th>Verbal–spatial recording(b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Attend auditory</td>
<td>0.60</td>
<td>0.18</td>
</tr>
<tr>
<td>Distraction</td>
<td>0.56</td>
<td>0.09</td>
</tr>
<tr>
<td>Attend verbal</td>
<td>0.66</td>
<td>0.11</td>
</tr>
<tr>
<td>Distraction</td>
<td>0.57</td>
<td>0.09</td>
</tr>
<tr>
<td>Ignore both</td>
<td>0.71</td>
<td>0.11</td>
</tr>
</tbody>
</table>

\(n = 11\). \(n = 13\).

ing that performance on the tone comparison task showed a greater decrease between a relatively brief (0.5-s) and relatively lengthy (10.0-s) ITI when the participants were required to perform a silent distractor task than when they were allowed to rehearse the standard tone. Both verbal and auditory distraction resulted in significantly lower levels of tone comparison performance at a 10.0-s ITI relative to the rehearsal condition.

In the present study, we also attempted to verify that the effect of rehearsal genuinely operates on a retention process, rather than simply affecting the perceptual processing of the stimuli to be compared. The inclusion of brief (0.5 s) ITIs was intended to allow the assessment of performance under different attentional sets in a situation where neither rehearsal nor distractor task activity could affect performance of the tone comparison task. Although the results of Experiment 1 indicated that performance at this ITI was not affected by the participants’ attentional set, Experiment 2 indicated that effects of attentional set on the processing of the stimuli must be considered in any delayed comparison task involving blocked instructional conditions. An effect of attentional set was found even at the brief ITI in this experiment, which suggests that the requirement to attend to a series of tones presented prior to the standard tone can reduce the perceptual processing of the standard tone. This finding is consistent with a storage-forgetting model of memory for pitch presented by Massaro (1970), in which auditory sensory memory strength is affected not only by the rate of forgetting but also by the amount of perceptual processing that the standard tone undergoes.

The effect of perceptual processing emerged in performance at the 10.0-s ITI as an effect of block type on performance on rehearsal trials. However, a difference in performance between rehearsal and distraction instructions within distractor blocks also was found. It could not be attributed to a difference in the perceptual processing of the standard tone, given that the participants did not know what the intone task would be until after the first tone was presented. Some form of tone rehearsal process must be invoked to explain this finding.

We do not, however, wish to suggest that the rehearsal of auditory sensory memory is fully analogous to the rehearsal of
verbal information. The relatively small magnitude of the distractor task effect found in the present study stands in sharp contrast to the robust effects that are obtained in tests of verbal memory. For example, the ability to retain three consonants is lost profoundly across 18 s filled with a backward counting task (Peterson & Peterson, 1959). Our data thus seem compatible with the conventional notion that the retention of verbal information is much more effective than the retention of fine-grained sensory information. However, the data also show that some form of rehearsal of this sensory information, presumably depending on auditory imagery, can occur.

It remains for future experiments to define the conditions under which rehearsal of tones can be useful to performance. Keller and Cowan (1994, Experiment 2) failed to find an effect of tone rehearsal in an experiment that differed from those described here in at least two important ways. First, it involved a same-different response rather than the present higher-lower response. Second, though the interfering task required imagination of a melody during the ITI, the melody to be imagined was a well-known one rather than a novel, span-length series presented immediately before the standard tone.

It is theoretically possible that performance of the present task involved a categorical recoding of the pitch of the tones into some form of verbal codes to be compared (e.g., "rather high," "sort of high," etc.). We cannot completely dismiss this possibility on the basis of the present data, but under direct questioning, none of the participants in Experiment 2 reported using such a strategy, and it seems unlikely that such a strategy was used, given the considerable memory loss that was observed across the ITI even in the absence of a distracting task. The adjustment of the tone discrimination difficulty to the participant's individual threshold and the use of a roving standard tone both were intended to make it difficult for participants to use a verbal code of this sort.

A final question addressed by the present study concerns the extent to which auditory images parallel auditory perceptions. The experiments conducted by Deutsch (1970) and Pechmann and Mohr (1992) both found that performance dropped from near perfect under a condition in which the ITI was silent and the participant was free to rehearse the tone to a chance level when six additional tones were presented during the interval, even though the participants were not required to attend to these interfering tones. In contrast, in each of our experiments, a much smaller difference in performance was obtained between a rehearsal condition and a distraction condition in which the participant recalled an average of five to six tones during a silent ITI. Imagining additional tones during the ITI apparently does not affect performance on a pitch comparison task to the same degree as actually perceiving additional tones during the ITI.

In neither of the present experiments was the auditory distractor task reliably more detrimental to tone comparison performance than a verbal distractor task. Both the Deutsch (1970) and Pechmann and Mohr (1992) studies indicated that the presentation of interfering tones degraded tone comparison performance much more than the presentation of interfering verbal auditory stimuli, regardless of whether the verbal items were attended or not. It therefore appears that interfering auditory stimulation results in specific interference with memory for pitch and thus depends on target-distractor similarity, in addition to any effect that such stimulation may have in reducing the attentional resources necessary for rehearsal. In contrast, the effect of interfering imagery during the ITI of a tone comparison task, in the present study, was comparable regardless of whether the imagery comprised tone or word sequences. Therefore, it probably can be said that tone sequence imagery affects two-tone discriminations solely by disrupting the rehearsal process rather than by directly degrading the auditory sensory representation. (The only qualification of that statement is that there conceivably also could have been some tone-specific imagery interference, which could have been obscured in the results if the small but significant advantage in distractor task accuracy for the verbal distraction condition, shown in Tables 1 and 2, arose from greater attention to the verbal distractor task than to the auditory distractor task.) An effect of tone rehearsal could operate by reactivating, enhancing, or supplementing a fading sensory memory representation.

References


Received June 7, 1989

Revision received July 29, 1994

Accepted August 1, 1994

---

**Research Awards in Experimental Psychology**

The Division of Experimental Psychology of the American Psychological Association (Division 3) announces a series of up to five annual research awards. These reviews are to be based on review of the research submitted to or published in the APA’s Journals of Experimental Psychology each year by relatively new investigators. The intention is to provide early recognition to new scholars whose research contributions are especially promising. These awards are

- **Division of Experimental Psychology (Annual)**
  - Young Investigator Award in Experimental Psychology: Animal Behavior Processes;

- **Division of Experimental Psychology (Annual)**
  - Young Investigator Award in Experimental Psychology: Human Perception and Performance;

- **Division of Experimental Psychology (Annual)**
  - Young Investigator Award in Experimental Psychology: Learning, Memory, and Cognition;

- **Division of Experimental Psychology (Annual)**
  - Young Investigator Award in Experimental Psychology: General;

and

- **Division of Experimental Psychology (Annual)**
  - Young Investigator Award in Experimental Psychology: Applied.

These awards will be announced and given to the winners each year at the Division 3’s business meeting held at the APA annual convention.