

## Minimal effects of visual memory training on auditory performance of adult cochlear implant users

Sandra I. Oba, MS;<sup>1\*</sup> John J. Galvin III, BA;<sup>1</sup> Qian-Jie Fu, PhD<sup>1–2</sup>

<sup>1</sup>Division of Communication and Auditory Neuroscience, House Research Institute, Los Angeles, CA; <sup>2</sup>Department of Biomedical Engineering and Neuroscience Program, University of Southern California, Los Angeles, CA

**Abstract**—Auditory training has been shown to significantly improve cochlear implant (CI) users' speech and music perception. However, it is unclear whether posttraining gains in performance were due to improved auditory perception or to generally improved attention, memory, and/or cognitive processing. In this study, speech and music perception, as well as auditory and visual memory, were assessed in 10 CI users before, during, and after training with a nonauditory task. A visual digit span (VDS) task was used for training, in which subjects recalled sequences of digits presented visually. After the VDS training, VDS performance significantly improved. However, there were no significant improvements for most auditory outcome measures (auditory digit span, phoneme recognition, sentence recognition in noise, digit recognition in noise), except for small (but significant) improvements in vocal emotion recognition and melodic contour identification. Posttraining gains were much smaller with the nonauditory VDS training than observed in previous auditory training studies with CI users. The results suggest that posttraining gains observed in previous studies were not solely attributable to improved attention or memory and were more likely due to improved auditory perception. The results also suggest that CI users may require targeted auditory training to improve speech and music perception.

**Key words:** auditory learning, auditory perception, auditory training, cochlear implant, digit span, hearing loss, memory training, music perception, rehabilitation, speech recognition.

### INTRODUCTION

While cochlear implants (CIs) have restored hearing to many deaf individuals, great variability remains in CI patient outcomes. Many CI patients perform well under optimal listening conditions (e.g., clear speech in quiet, simple music perception), but some do not, even after years of experience with their device. Even top-performing CI patients have great difficulty with challenging listening conditions (e.g., speech in noise, complex music perception). These deficits are likely due to the limited spectrotemporal resolution of the device and/or the implanted ear [1]. Recent advances in CI technology (e.g., high stimulation rates, current steering) have shown only small and/or inconsistent improvements [2–5]. In contrast, auditory training has been shown to improve speech and music performance for CI patients with years of experience with their device [6–9]. These results suggest that postlingually

**Abbreviations:** ADS = auditory digit span, CI = cochlear implant, HA = hearing aid, HINT = Hearing in Noise Test, MCI = melodic contour identification, NH = normal hearing, RM ANOVA = repeated measures analysis of variance, RMS = root-mean-square, SNR = signal-to-noise ratio, SRT = speech recognition threshold, VDS = visual digit span.

\*Address all correspondence to Sandra I. Oba, MS; Division of Communication and Auditory Neuroscience, House Research Institute, 2100 West Third St, Los Angeles, CA 90057; 213-353-7054; fax: 213-483-8789.

Email: [soba@hei.org](mailto:soba@hei.org)

<http://dx.doi.org/10.1682/JRRD.2011.12.0229>

deafened CI patients may not have learned to use all of the information provided by their device.

Auditory training has been shown to improve CI users' speech understanding in quiet [6–7,10–11] and in noise [9], as well as music perception [8,12–14]. Auditory training has also been shown to improve speech performance in normal-hearing (NH) subjects listening to acoustic CI simulations [15–22]. While the greatest post-training gains in performance were often for the trained task, training benefits in these studies often generalized to untrained tasks and/or conditions, e.g., from monosyllable word training to recognition of phonemes and words in sentences [7] or from melodic contour training to familiar melody recognition [8]. Interestingly, for NH subjects listening to CI simulations, training identification of environmental sounds improved speech recognition, while speech training did not improve recognition of environmental sounds [17]. These studies and others reveal a variety of outcome patterns for different training tasks, protocols, and subject groups that may be highly variable, making it difficult to predict generalization effects [23].

Recent studies suggest that the benefits of auditory training may not solely be due to improved auditory perception. Cognitive factors, such as attention and memory, may help listeners to “tune in” to relevant acoustic features [24–26]. Training with “identical” stimuli (i.e., stimuli with the same frequency and amplitude) has been shown to improve NH listeners' frequency discrimination [24]. In the study, the training feedback was random and unrelated to the stimuli; subjects were led to believe that there was a difference in the tones, which may have increased attention to details within the stimuli. In the same study, training with an unrelated visual task (a silent visuospatial puzzle video game) also improved frequency discrimination, suggesting that general attention and/or state of arousal may play a fundamental role in auditory learning. NH musicians have shown better perceptual thresholds than nonmusicians for frequency discrimination, backward masking, and an auditory attention task measuring reaction time [26]; the authors suggest that long-term musical experience may strengthen cognitive function, which in turn may benefit auditory perception.

Performance on memory tasks has been significantly correlated with speech recognition performance in children with CIs [27–28], even after several years of experience with their CIs [29]. Pediatric CI users' auditory digit span (ADS)—a commonly used measure of memory

capacity—has been significantly correlated with speech perception, speech production, language comprehension, and reading [27]. Strong correlations between speaking rate and forward and backward digit span have also been reported for pediatric CI users, suggesting that processing of immediate memory may limit speech understanding [28]. Pediatric CI users have been shown to perform more poorly than their NH peers on a memory task even when the stimuli were presented visually and did not require a verbal response [30]. Training pediatric CI patients using a program of auditory, visuospatial, and combined auditory/visuospatial computer tasks significantly improved verbal and nonverbal memory tests scores, as well as sentence repetition skills. In the video game-like exercises, the level of difficulty was adapted in terms of span length and complexity according to subject performance [31].

In previous auditory training studies with CI users [6–9], it was unclear whether improvements were primarily due to improved auditory perception or cognitive processing (e.g., attention, memory, concentration). It is possible that training with a nonauditory task that targets cognitive processing might yield similar gains in auditory performance. In this study, a nonauditory training task intended to engage attention and improve memory was used to train CI subjects. For the purposes of this article, the general terms “attention” and “memory” are used to denote cognitive processes inherent to information processing and likely interconnected in regards to learning. Speech and music perception were assessed in 10 CI users before, during, and after training with a forward visual digit span (VDS) task in which listeners were asked to recall visually presented sequences of digits. Auditory and visual memory was also tested before, during, and after the VDS training. Given the correlation between digit spans scores and speech performance [27–29] and the significant auditory benefits from nonauditory training [24], we hypothesized that the VDS training would improve CI users' auditory perceptual performance.

## METHODS

### Subjects

Ten adult, postlingually deafened CI users participated in this study. Relevant demographic information is shown in the **Table**. All subjects were required to be native speakers of American English and to have at least

**Table 1.**

Cochlear implant (CI) subject demographics.

Subject	Sex	Age (yr)	Etiology	CI Use	Device	Processor
1	M	24	Meningitis	6 mo	Nucleus Freedom (R); Hearing Aid (L)	CP810
2	M	60	Unknown	10 yr	Advanced Bionics 90k/ HiFocus (L)	Harmony
3	F	25	Unknown	3 yr	Nucleus Freedom (R)	Freedom
4	M	62	Meningitis	19 yr	Nucleus 22 (R)	ESPrIt 3G
5	F	91	Unknown	6 mo (R); 4 yr (L)	Nucleus 5 (R); Nucleus Freedom (L)	CP810 (R); Freedom (L)
6	M	20	Genetic	16 yr	Nucleus 22 (L)	Freedom
7	F	53	Unknown	1 yr (R); 4 yr (L)	Advanced Bionics 90k/ HiFocus(R); Advanced Bionics 90k/ HiFocus (L)	Harmony (R); Harmony (L)
8	F	64	Meniere's Disease/Otosclerosis	4 yr	Nucleus Freedom (L)	Freedom
9	M	72	Unknown/Noise Exposure	8 mo	Nucleus Freedom (R); Hearing Aid (L)	Freedom
10	F	50	Unknown/Ototoxicity	16 yr (R) 4 yr (L)	Advanced Bionics Clarion (R); Advanced Bionics HiRes 90k(L)	BTE II (R); Harmony (L)

F = female, L = left, M = male, R = right.

6 months of experience with their implant device. None of the subjects had participated in any previous auditory training experiments. Five subjects (S1, S2, S3, S7, and S10) had some experience with speech recognition experiments, while the other five (S4, S5, S6, S8, and S9) had no prior experience. All provided informed consent before participating and all were reimbursed for their time and expenses associated with testing in the laboratory and training at home.

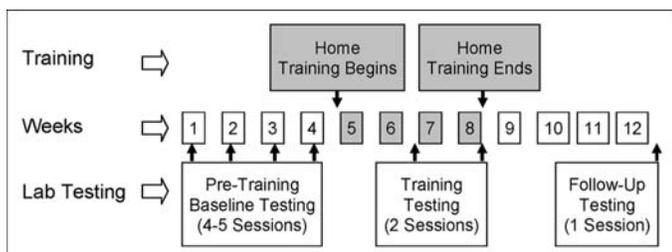
Subjects were tested using their clinical speech processors set to their preferred “everyday” program and volume control settings. If a subject regularly wore a hearing aid (HA) in the contralateral ear in combination with the CI, testing and training was performed with the CI and HA using the everyday, clinical settings. Subjects were instructed to use the same CI (and HA) settings for all testing and training sessions.

### General Testing and Training Timeline

Because intersubject variability can be quite large in CI studies, it can be difficult to separate within-subject training effects from across-subject variability. It is also difficult to establish appropriate CI control groups of sufficient numbers (i.e., balanced in terms of subject age, duration of deafness, device type, baseline performance), especially for training studies that require an extended

time commitment. In this study, a within-subject control procedure was used instead of an across-subject control group, with each subject serving as his or her own control. The within-subject control procedure required extensive baseline performance measures before training began. Baseline performance was repeatedly measured in the laboratory once per week for a minimum of four sessions. At the fourth session, if performance improved by more than the standard deviation of the first three sessions, baseline performance was measured a fifth time. The results of the last two test sessions were averaged and considered the pretraining baseline scores.

After completing pretraining baseline measures, subjects trained at home for a total of 10 h on their personal computers or on loaner laptops using custom software (“Sound Express,” developed by Qian-Jie Fu). Participants were instructed to train for ~30 min per day, 5 days a week, for 4 weeks. Posttraining performance for all tests was measured after completing 5 and 10 h of training. Training was stopped after 10 h, and participants returned to the laboratory 1 month later for follow-up measures to see whether any posttraining gains in performance were retained. **Figure 1** illustrates the testing and training schedule.



**Figure 1.**  
Schematic of experiment testing and training schedule.

### Test Methods and Materials

Assessment measures covered a broad range of auditory and memory/attention tasks. Memory was assessed with (1) VDS and (2) ADS. Speech recognition in noise was assessed with (1) the Hearing in Noise Test (HINT) sentences in steady noise, (2) HINT sentences in babble, (3) digits in steady noise, and (4) digits in babble. Speech recognition in quiet was assessed with (1) vowel identification and (2) consonant identification. Other auditory tests included (1) vocal emotion recognition and (2) melodic contour identification (MCI).

VDS was measured using an adaptive (1-up/1-down) procedure. Stimuli included digits 0 through 9. During testing, digits were randomly selected and presented visually in sequence on the computer screen (e.g., “8–3–0”). Each digit was individually shown on the computer screen for approximately 0.5 s. Subjects responded by clicking on response boxes (labeled “0” through “9”) shown on a computer screen or by typing in the numbers on the keyboard. The initial sequence contained three digits. Depending on the correctness of response, the number of digits presented was either increased or decreased (two-digit step size for the first two reversals and one-digit step size for subsequent reversals). Each test run was 25 trials. The VDS score represented the number of digits that could be correctly recalled on 50 percent of the trials and was calculated as the average number of digits correctly recalled across all but the first two reversals. One test run was measured at each test session.

Similar to VDS, ADS tested the recall of a sequence of digits immediately after auditory presentation. The adaptive procedure and method of scoring were the same as for the VDS task, except that the sequence of digits was presented in an auditory-only context (no visual cues). One test run of 25 trials was measured at each test session.

Speech recognition thresholds (SRTs) for HINT sentences [32] were measured using an open-set, adaptive (1-up/1-down) procedure [33]. The SRT was defined as the signal-to-noise ratio (SNR) that produced 50 percent correct whole sentence recognition. HINT SRTs were measured in steady speech-shaped noise (1,000 Hz cutoff frequency, –12 dB/octave) and in six-talker speech babble. The SNR was calculated according to the long-term root-mean-square (RMS) of the speech and noise stimuli. HINT stimuli included 260 sentences produced by one male talker and were of easy to moderate difficulty in terms of vocabulary and syntactic complexity (e.g., “The picture came from a book.”). During each test run, a sentence was randomly selected (without replacement) from the 260-sentence stimulus set and presented in noise. Subjects were asked to repeat what they heard as accurately as possible. If they repeated the entire sentence correctly, the SNR was reduced by 2 dB; if not, the SNR was increased by 2 dB. Subjects were allowed to hear one repetition of the test sentence. On average, 15 sentences were presented within each test run, and the mean of the final six reversals in SNR was recorded as the SRT. Two runs each in steady noise and speech babble were measured during each test session. Individual sentences were not repeated within a given test run; however, some sentences may have been repeated across runs.

Digit in noise SRTs were measured using a closed-set, adaptive (1-up/1-down) procedure, converging on the SNR that produced 50 percent correct sequence identification [33]. Stimuli included digits 0 through 9 produced by one male talker. During testing, three digits were randomly selected and presented in sequence (e.g., “five-one-nine”) in the presence of background noise. Subjects responded by clicking on response boxes (labeled “zero” through “nine”) on a computer screen or by typing the numbers on a keyboard. Subjects were allowed to repeat the stimulus up to three times. If the participant correctly recognized the entire sequence, the SNR was reduced by 2 dB; if not, the SNR was increased by 2 dB. Each test run was 25 trials. All reversals in SNR were used to calculate the digits in noise SRT. Two runs each in steady noise and speech babble were measured at each test session.

Vowel recognition was assessed with 12 vowels, presented in /h/-vowel-/d/ context (heed, hid, head, had, who’d, hood, hod, hud, hawed, heard, hoed, hayed). Vowel stimuli were digitized natural productions from five male and five female talkers [34]. A stimulus token was randomly chosen for each trial without replacement

from the vowel set and presented to the listener. Subjects responded by clicking on one of 12 response choices on a computer screen; the response buttons were labeled orthographically. Subjects were allowed to repeat each stimulus up to three times. Results were reported as percentage correct. Each test block included 120 trials (12 vowels  $\times$  10 talkers), and one block was completed at each test session.

Consonant recognition was measured with 20 nonsense syllables presented in /a/-consonant-/a/ context (/ b d g p t k m n l r y w f s j v z ð tʃ dʒ /). Consonant stimuli were digitized natural productions from five male and five female talkers [35]. For each trial, a stimulus token was randomly chosen without replacement from the consonant set and presented to the listener. Subjects responded by clicking one of 20 response choices displayed on a computer screen; the response buttons were labeled orthographically. Listeners were allowed to repeat each stimulus up to three times. Results were reported as percent correct. Each test block included 200 trials (20 vowels  $\times$  10 talkers), and one block was completed at each test session.

Vocal emotion recognition was assessed using stimuli from the House Ear Institute Emotional Speech Database [36]. Stimuli consisted of 10 semantically neutral sentences (e.g., “The year is almost over.”) produced by one male and one female talker according to five target emotions: angry, anxious, happy, sad, and neutral. In each trial, a sentence was randomly selected without replacement from the stimulus set and presented to the listener. Subjects responded by clicking on one of the five response choices shown on a computer screen (labeled “angry,” “anxious,” “happy,” “sad,” and “neutral”). Subjects were allowed to repeat each stimulus up to three times. Responses were scored in terms of percent correct. Each test block included 100 trials (10 sentences  $\times$  5 emotions  $\times$  2 talkers), and one block was completed at each test session.

MCI was used to assess musical pitch perception [8]. Stimuli included nine 5-note melodic contours that represented simple changes in pitch (“rising,” “falling,” “flat,” “flat-rising,” “flat-falling,” “rising-flat,” “falling-flat,” “rising-falling,” and “falling-rising”). The spacing between successive notes in each contour was 1, 2, or 3 semitones, and the lowest note in the contour was A3 (220 Hz). The note duration was 250 ms and the interval between notes was 50 ms. The contours were played by a piano sample. In each trial, a contour was randomly

selected from the stimulus set and presented to the listener. Subjects responded by clicking on one of nine choices shown on a computer screen, which were visually labeled according to the pitch patterns. Listeners were allowed to repeat each stimulus up to three times. Responses were scored in terms of percent correct. Each test block included 54 trials (9 contours  $\times$  3 semitone spacings  $\times$  2 repeats of stimuli) and one block was completed at each test session.

Testing was conducted in sound field in a sound-treated booth (IAC; New York, New York). All auditory stimuli were delivered via single loudspeaker (Reveal, Tannoy; Coatbridge, Scotland) at a fixed output of 65 dBA. For tests involving speech in noise, the target SNR was calculated according to the long-term RMS of the speech and noise; the level of the combined speech and noise was then adjusted to achieve the target output (65 dBA). Participants were seated directly facing the loudspeaker. No training or trial-by-trial feedback was provided during test sessions. The same performance measures and methods were used for pretraining baseline, posttraining, and follow-up test sessions.

### Training Methods and Materials

VDS training began after completing pretraining baseline measures. Training was conducted at home using custom software loaded onto each subjects’ personal computers or loaner laptops. Subjects were instructed on how to install and use the software and to immediately contact the research staff if there were any problems with the home training. The software allowed for remote monitoring of training sessions, letting researchers view subjects’ training frequency, duration, and performance. Subjects were asked to train for 30 min/d, 5 d/wk, for 4 wk, for a total of 10 training hours.

The VDS training and VDS testing stimuli and procedure were the same except that VDS training included trial-by-trial visual feedback. If the subject responded correctly, visual feedback was provided, a new sequence was presented, and the number of digits in the sequence was increased. If the subject responded incorrectly, visual feedback was provided, in which the incorrect response and the correct response were shown on the computer screen. Then the next digit sequence was presented and the number of digits in the sequence was decreased. The step size was two digits for the first two reversals and one digit for subsequent reversals. Each training exercise

consisted of 25 digit sequences and took 10 to 15 min to complete.

## RESULTS

### Home Training

All subjects completed the VDS home training; the mean training time was 613 min (range: 580–665 min). Home training data (rather than laboratory testing data) showed that the mean VDS score improved from 6.94 at the start of training to 7.53 and 8.31 after 5 and 10 h of training, respectively. A one-way repeated measures analysis of variance (RM ANOVA) showed a significant effect for home training session ( $F_{2,18} = 17.41, p < 0.001$ ). Post hoc Bonferroni *t*-tests showed significant differences between start of training and 10 h of training ( $p < 0.001$ ), and between 5 hours and 10 h of training ( $p = 0.01$ ).

### Laboratory Testing

**Figure 2** shows baseline (pretraining) performance for all test measures across test sessions. Within each panel, the solid line shows the linear regression fit across all subject data. The slopes of the regressions were quite shallow, and  $p > 0.05$  in all cases. The mean of the last two sessions was considered the baseline score for each test (shaded region in each panel).

**Figure 3** shows individual and mean VDS scores at baseline, 5 and 10 h posttraining, and follow-up. Although there was considerable intersubject variability, most subjects exhibited some improvement after training. The mean VDS score improved from 6.72 at baseline to 7.77 and 7.97 after 5 and 10 h of training, respectively; follow-up scores were largely unchanged (7.83 digits) from posttraining performance. A one-way RM ANOVA showed that VDS scores differed significantly across the test sessions ( $F_{3,27} = 9.73, p < 0.001$ ). Post hoc Bonferroni *t*-tests showed significant differences between baseline and 5 h posttraining ( $p = 0.002$ ), 10 h posttraining ( $p < 0.001$ ), and follow-up ( $p = 0.001$ ).

**Figure 4** shows mean performance for all test measures at baseline, 5 and 10 h posttraining, and follow-up. For most auditory measures, posttraining improvements were quite small. One-way RM ANOVAs showed no significant effect of test sessions for ADS, HINT in steady noise or babble, digits in steady noise or babble, vowels, or consonants ( $p$ -values ranged from 0.07 to 0.88). However, one-way RM ANOVAs showed significant effects

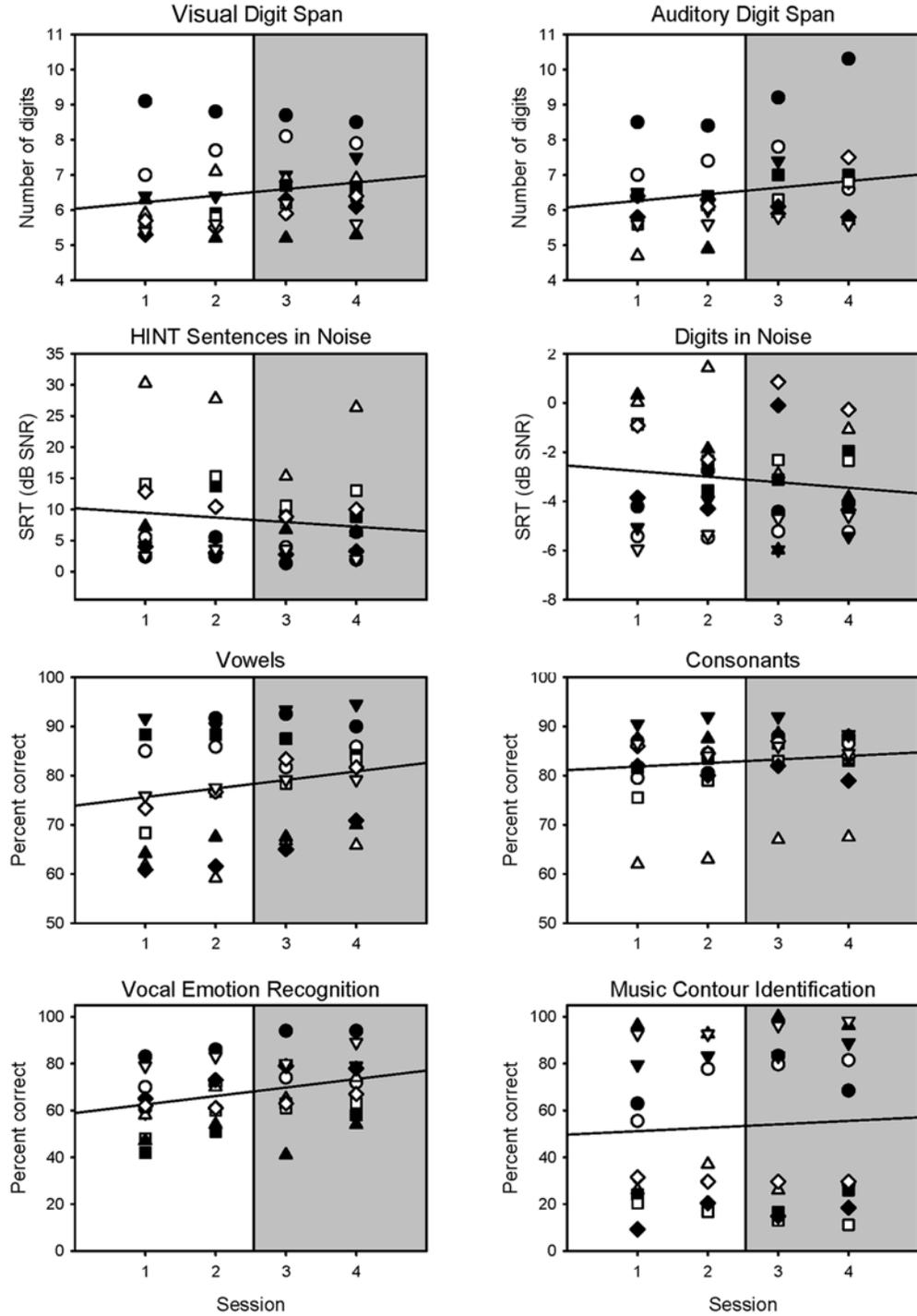
for test session for vocal emotion recognition ( $F_{3,27} = 3.31, p = 0.03$ ) and MCI ( $F_{3,27} = 2.98, p = 0.049$ ). Post hoc Bonferroni *t*-tests did not reveal significant effects; however, the power of the performed test was low (0.511 for vocal emotion identification and 0.444 for MCI). Note that the mean improvement after 10 h of training was only 3.45 and 4.35 percentage points for vocal emotion recognition and MCI, respectively.

**Figure 5** compares the mean posttraining gain for various outcome measures after ~10 h of VDS training in the present study or after ~10 h of auditory training in previous CI studies [7–9]. For all outcome measures, the posttraining gains were larger with auditory than with the present nonauditory training. Posttraining gains for speech understanding in noise (**Figure 5(a)**) were 0.07 to 1.11 dB with the present VDS training, versus 0.49 to 3.82 dB with the previous digits in babble auditory training [9]. Posttraining gains for phoneme recognition (**Figure 5(b)**) were 0.81 to 1.07 percentage points with the present VDS training versus 13.54 to 15.81 points with the previous phonemic contrast auditory training [7]. The mean posttraining gain for MCI (**Figure 5(b)**) was 3.45 points with the present VDS training versus 28.30 percentage points with the previous MCI auditory training [8].

Subject demographics were compared with pre- and posttraining performance. For pretraining performance, significant ( $p < 0.05$ ) correlations (Spearman rank order) were found between age at testing and vocal emotion recognition and between CI experience and consonant recognition. For posttraining performance, significant correlations were found between age at testing and digits in babble, between age at testing and vocal emotion recognition, and between CI experience and HINT in steady noise.

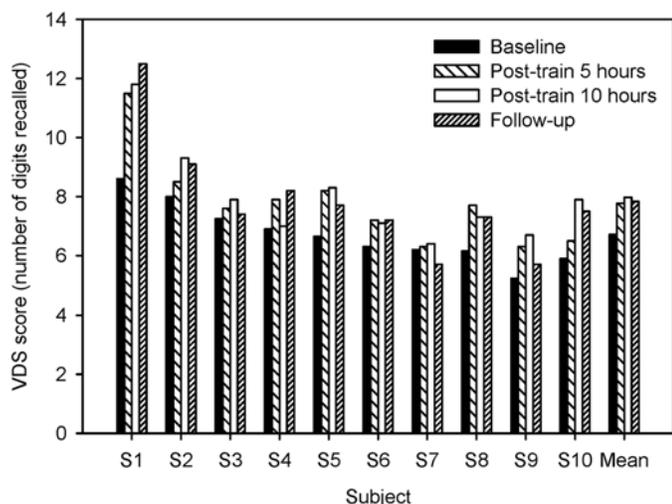
## DISCUSSION

While the present VDS training significantly improved VDS scores, it provided little to no benefit for auditory perceptual measures. There were no significant improvements in ADS, HINT sentence recognition in noise, digits recognition in noise, or phoneme recognition in quiet, and only small (but significant) improvements observed for vocal emotion recognition and MCI. The results do not support the hypothesis that nonauditory training can improve auditory perception. Further, the results suggest that the posttraining gains observed in previous CI training studies [6–10] were not solely due to improved attention, memory,



**Figure 2.**

Baseline (pretraining) performance for all test measures as function of test session. Final four sessions are shown; in cases where five baseline sessions were required to achieve asymptotic performance, only data from last four sessions are included. Different symbols show individual subject data. For Hearing in Noise Test (HINT) Sentences and Digits in Noise, mean performance across steady noise and babble conditions is shown. Solid lines show linear regressions fit across all subject data. Shaded area shows scores used to represent baseline performance. SNR = signal-to-noise ratio, SRT = speech recognition threshold.



**Figure 3.** Visual digit span (VDS) score for individual subjects.

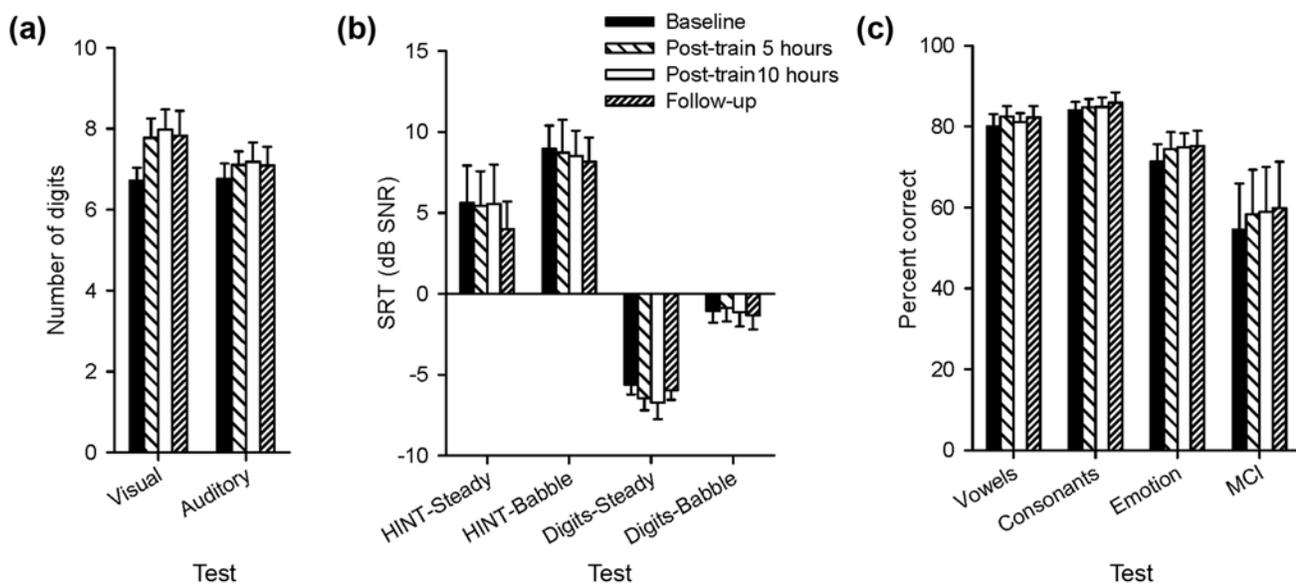
and/or cognitive processes, but rather to improved auditory perception via auditory training.

We should note several potential limitations to the present study. First, the small number of participants ( $n = 10$ ) precludes any strong conclusions being drawn from the present results. While there were few instances of sig-

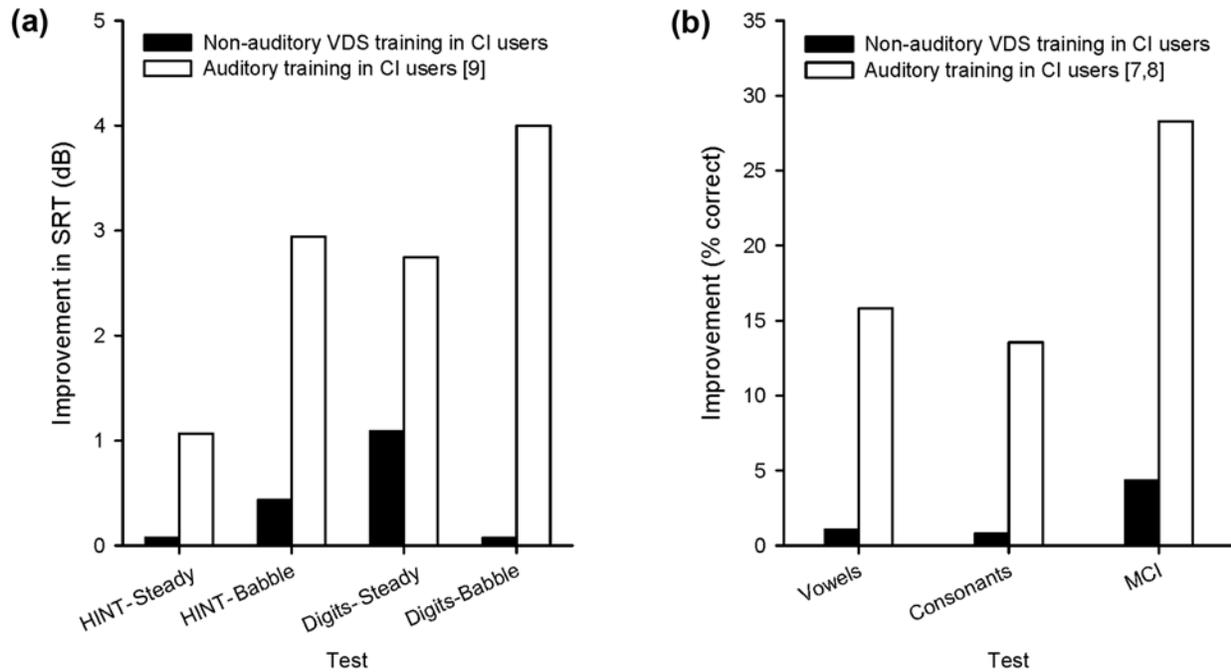
nificant improvement after VDS training, the power was low, increasing the likelihood of type II error. A greater number of subjects would improve power and might reveal significant benefits with VDS training. However, we expect that posttraining gains would remain smaller than with auditory training.

Second, VDS may not have been the most effective nonauditory training task. Multiple nonauditory training tasks (e.g., VDS, color sequence identification, word list identification) might have been more beneficial than the single task used in this study. Training that incorporated both auditory and visual tasks might have yielded even further benefits, although it would have been difficult to ascertain exactly what was being learned. Note that non-auditory training (i.e., learning to play a silent video game) has been shown to produce small but significant improvements in auditory perception [24]; however, in that same study, auditory training or even exposure to auditory stimuli (passive learning) provided much larger performance gains.

Third, because the VDS training was not necessarily “enjoyable,” it may not have effectively engaged subjects’ attention. Subjects reported that the VDS task was difficult and that they often broke up the daily training into a couple of sessions. However, similar improve-



**Figure 4.** Mean performance (across subjects) for (a) visual and auditory digit span; (b) speech recognition in noise; and (c) phoneme recognition, vocal emotion recognition, and melodic contour identification (MCI) in quiet. Error bars show 1 standard error. HINT = Hearing in Noise Test, SNR = signal-to-noise ratio, SRT = speech recognition threshold.



**Figure 5.**

Mean posttraining performance gains for cochlear implant (CI) users after present visual digit span (VDS) training (black bars) or after auditory training in previous studies (white bars). Posttraining gains are shown for speech recognition (a) in noise and (b) for speech and melodic contour identification (MCI) in quiet. HINT = Hearing in Noise Test, SRT = speech recognition threshold.

ments in VDS were observed across home training (6.94 to 7.85 digits) and laboratory testing (6.72 to 7.97 digits) during the 10 h of training, suggesting active engagement with the VDS training.

Finally, this study did not have an independent control group and used a within-subject control procedure. The extensive baseline testing prior to training was intended to minimize procedural learning effects [37]. However, because of the repeated testing (baseline, post-training, and follow-up sessions), it is possible that some familiarization and/or memorization may have affected closed-set outcome measures (i.e., phoneme recognition, vocal emotion recognition, and MCI). Note that the same within-subject control and extensive baseline testing procedures were used in our previous auditory training studies [6–9]. In those studies, much larger posttraining gains occurred after auditory training with novel stimuli (not used for testing) than with the present nonauditory VDS training. This suggests that the repeated exposure may not have been a limiting factor for this study.

In previous studies with pediatric CI patients [27–28], significant correlations were observed between ADS and closed-set word recognition, open-set word recogni-

tion, and recognition of words in sentences. In this study, both before and after training, there were significant correlations only between VDS and ADS and between ADS and vowel recognition; no other correlations were observed. The shorter digit spans and poorer speech performance in the previous pediatric CI studies [27–28] may reflect developmental effects and/or memory processing problems that were not present in the present adult subjects.

There has been much recent research interest in “brain-training” computer games and software designed to reduce cognitive decline in older adults. Some studies show generalized improvements beyond the trained task [38–39], while others do not [40–41]. Computer exercises (using auditory stimuli) have been proposed to improve auditory information processing in adults aged 65 and older [39]. In these exercises, the level of difficulty is adapted according to subject performance; exercises include discrimination of easily confused syllables and recall of details in a verbally presented story. Such training has shown significant performance gains for the trained tasks, as well as improvements in standardized measures of cognitive function (e.g., backward digit span

and delayed recall of word lists) that were not directly trained. While the present study used a single visual training task, it is possible that an auditory memory/attention task or a battery of training exercises unrelated to the present outcome measures may be more beneficial.

## CONCLUSIONS

The present nonauditory VDS training provided only minimal gains in auditory perception tasks. While memory, attention, and cognitive processes are part of auditory learning, the very small posttraining gains with the VDS training suggest that auditory training may be required to improve auditory perception and that the positive training outcomes in previous studies [6–9] were not solely attributable to improvements in cognitive processes such as attention and memory. CI users may require specific targeted auditory tasks for training to improve speech recognition and auditory perception.

## ACKNOWLEDGMENTS

### Author Contributions:

*Study concept and design:* Q. J. Fu.

*Acquisition of data:* S. I. Oba.

*Analysis and interpretation of data:* S. I. Oba, J. J. Galvin III, Q. J. Fu.

*Drafting of manuscript:* S. I. Oba, J. J. Galvin III, Q. J. Fu.

*Critical revision of manuscript for important intellectual content:*

S. I. Oba, J. J. Galvin III, Q. J. Fu.

*Statistical analysis:* S. I. Oba, J. J. Galvin III.

*Obtained funding and study supervision:* Q. J. Fu.

**Financial Disclosures:** The authors have declared that no competing interests exist.

**Funding/Support:** This material is based on work supported by the National Institutes of Health National Institute on Deafness and Other Communication Disorders (grant RO1-DC004792).

**Additional Contributions:** The authors thank all of the CI users who participated in this study for their tireless work and support of CI research.

**Institutional Review:** All subjects provided informed consent before participation and all study-related procedures were approved by the institutional review board of St. Vincent Medical Center, Los Angeles.

**Participant Follow-up:** Subjects have been informed of the intent to publish and that they may request details of publication from the authors.

## REFERENCES

- Shannon RV, Fu QJ, Galvin JJ 3rd. The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryngol Suppl.* 2004;552(552):50–54. [PMID:15219048]
- Donaldson GS, Dawson PK, Borden LZ. Within-subjects comparison of the HiRes and Fidelity120 speech processing strategies: speech perception and its relation to place-pitch sensitivity. *Ear Hear.* 2011;32(2):238–50. [PMID:21084987]  
<http://dx.doi.org/10.1097/AUD.0b013e3181fb8390>
- Firszt JB, Holden LK, Reeder RM, Skinner MW. Speech recognition in cochlear implant recipients: comparison of standard HiRes and HiRes 120 sound processing. *Otol Neurotol.* 2009;30(2):146–52. [PMID:19106769]  
<http://dx.doi.org/10.1097/MAO.0b013e3181924ff8>
- Plant K, Holden L, Skinner M, Arcaroli J, Whitford L, Law M-A, Nel E. Clinical evaluation of higher stimulation rates in the nucleus research platform 8 system. *Ear Hear.* 2007;28(3):381–93. [PMID:17485987]  
<http://dx.doi.org/10.1097/AUD.0b013e31804793ac>
- Vandali AE, Whitford LA, Plant KL, Clark GM. Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system. *Ear Hear.* 2000;21(6):608–24. [PMID:11132787]  
<http://dx.doi.org/10.1097/00003446-200012000-00008>
- Fu QJ, Galvin JJ 3rd. Maximizing cochlear implant patients' performance with advanced speech training procedures. *Hear Res.* 2008;242(1–2):198–208. [PMID:18295992]  
<http://dx.doi.org/10.1016/j.heares.2007.11.010>
- Fu QJ, Galvin JJ III, Wang X, Nogaki G. Moderate auditory training can improve speech performance of adult cochlear implant patients. *Acoust Res Lett Online.* 2005;6(3):106–11. <http://dx.doi.org/10.1121/1.1898345>
- Galvin JJ 3rd, Fu Q-J, Nogaki G. Melodic contour identification by cochlear implant listeners. *Ear Hear.* 2007;28(3):302–19. [PMID:17485980]  
<http://dx.doi.org/10.1097/01.aud.0000261689.35445.20>
- Oba SI, Fu Q-J, Galvin JJ 3rd. Digit training in noise can improve cochlear implant users' speech understanding in noise. *Ear Hear.* 2011;32(5):573–81. [PMID:21389857]  
<http://dx.doi.org/10.1097/AUD.0b013e31820fc821>
- Stacey PC, Raine CH, O'Donoghue GM, Tapper L, Twomey T, Summerfield AQ. Effectiveness of computer-based auditory training for adult users of cochlear implants. *Int J Audiol.* 2010;49(5):347–56. [PMID:20380610]  
<http://dx.doi.org/10.3109/14992020903397838>
- Wu JL, Yang HM, Lin YH, Fu QJ. Effects of computer-assisted speech training on Mandarin-speaking hearing-impaired children. *Audiol Neurotol.* 2007;12(5):307–12. [PMID:17536199]  
<http://dx.doi.org/10.1159/000103211>
- Gfeller K, Witt SW, Kim K-H, Adamek M, Coffman D. Preliminary report of a computerized music training pro-

- gram for adult cochlear implant recipients. *J Acad Rehabil Audiol.* 1999;32:11–27.
13. Gfeller K, Witt S, Stordahl J, Mehr M. The effects of training on melody recognition and appraisal by adult cochlear implant recipients. *J Acad Rehabil Audiol.* 2000;33:115–38.
  14. Yucel E, Sennaroglu G, Belgin E. The family oriented musical training for children with cochlear implants: speech and musical perception results of two year follow-up. *Int J Pediatr Otorhinolaryngol.* 2009;73(7):1043–52. [PMID:19411117] <http://dx.doi.org/10.1016/j.ijporl.2009.04.009>
  15. Davis MH, Johnsruide IS, Hervais-Adelman A, Taylor K, McGettigan C. Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *J Exp Psychol Gen.* 2005; 134(2):222–41. [PMID:15869347] <http://dx.doi.org/10.1037/0096-3445.134.2.222>
  16. Fu QJ, Galvin JJ 3rd. The effects of short-term training for spectrally mismatched noise-band speech. *J Acoust Soc Am.* 2003;113(2):1065–72. [PMID:12597199] <http://dx.doi.org/10.1121/1.1537708>
  17. Loebach JL, Pisoni DB. Perceptual learning of spectrally degraded speech and environmental sounds. *J Acoust Soc Am.* 2008;123(2):1126–39. [PMID:18247913] <http://dx.doi.org/10.1121/1.2823453>
  18. Loebach JL, Pisoni DB, Svirsky MA. Transfer of auditory perceptual learning with spectrally reduced speech to speech and nonspeech tasks: implications for cochlear implants. *Ear Hear.* 2009;30(6):662–74. [PMID:19773659] <http://dx.doi.org/10.1097/AUD.0b013e3181b9c92d>
  19. Nogaki G, Fu QJ, Galvin JJ 3rd. Effect of training rate on recognition of spectrally shifted speech. *Ear Hear.* 2007; 28(2):132–40. [PMID:17496666] <http://dx.doi.org/10.1097/AUD.0b013e3180312669>
  20. Rosen S, Faulkner A, Wilkinson L. Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants. *J Acoust Soc Am.* 1999;106(6): 3629–36. [PMID:10615701] <http://dx.doi.org/10.1121/1.428215>
  21. Stacey PC, Summerfield AQ. Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech. *J Acoust Soc Am.* 2007;121(5 Pt 1): 2923–35. [PMID:17550190] <http://dx.doi.org/10.1121/1.2713668>
  22. Stacey PC, Summerfield AQ. Comparison of word-, sentence-, and phoneme-based training strategies in improving the perception of spectrally distorted speech. *J Speech Lang Hear Res.* 2008;51(2):526–38. [PMID:18367694] [http://dx.doi.org/10.1044/1092-4388\(2008/038\)](http://dx.doi.org/10.1044/1092-4388(2008/038))
  23. Wright BA, Zhang Y. A review of the generalization of auditory learning. *Phil Trans R Soc B.* 2009;364(1515):301–11. [PMID:18977731] <http://dx.doi.org/10.1098/rstb.2008.0262>
  24. Amitay S, Irwin A, Moore DR. Discrimination learning induced by training with identical stimuli. *Nat Neurosci.* 2006;9(11):1446–48. [PMID:17028582] <http://dx.doi.org/10.1038/nn1787>
  25. Moore DR, Halliday LF, Amitay S. Use of auditory learning to manage listening problems in children. *Phil Trans R Soc B.* 2009;364(1515):409–20. [PMID:18986969] <http://dx.doi.org/10.1098/rstb.2008.0187>
  26. Strait DL, Kraus N, Parbery-Clark A, Ashley R. Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hear Res.* 2010;261(1–2):22–29. [PMID:20018234] <http://dx.doi.org/10.1016/j.heares.2009.12.021>
  27. Pisoni DB, Geers AE. Working memory in deaf children with cochlear implants: correlations between digit span and measures of spoken language processing. *Ann Otol Rhinol Laryngol Suppl.* 2000;185:92–93. [PMID:11141023]
  28. Pisoni DB, Cleary M. Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear Hear.* 2003;24(1, Suppl):106S–20S. [PMID:12612485] <http://dx.doi.org/10.1097/01.AUD.0000051692.05140.8E>
  29. Pisoni DB, Kronenberger WG, Roman AS, Geers AE. Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation. *Ear Hear.* 2011;32(1, Suppl):60S–74S. [PMID:21832890] <http://dx.doi.org/10.1097/AUD.0b013e3181ff58e>
  30. Cleary M, Pisoni DB, Geers AE. Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants. *Ear Hear.* 2001;22(5):395–411. [PMID:11605947] <http://dx.doi.org/10.1097/00003446-200110000-00004>
  31. Kronenberger WG, Pisoni DB, Henning SC, Colson BG, Hazzard LM. Working memory training for children with cochlear implants: a pilot study. *J Speech Lang Hear Res.* 2011;54(4):1182–96. [PMID:21173394] [http://dx.doi.org/10.1044/1092-4388\(2010/10-0119\)](http://dx.doi.org/10.1044/1092-4388(2010/10-0119))
  32. Nilsson M, Soli SD, Sullivan JA. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am.* 1994;95(2):1085–99. [PMID:8132902] <http://dx.doi.org/10.1121/1.408469>
  33. Levitt H. Transformed up-down methods in psychoacoustics. *J Acoust Soc Am.* 1971;49(2):2, 467. [PMID:5541744] <http://dx.doi.org/10.1121/1.1912375>
  34. Hillenbrand J, Getty LA, Clark MJ, Wheeler K. Acoustic characteristics of American English vowels. *J Acoust Soc Am.* 1995;97(5 Pt 1):3099–3111. [PMID:7759650] <http://dx.doi.org/10.1121/1.411872>

35. Shannon RV, Jensvold A, Padilla M, Robert ME, Wang X. Consonant recordings for speech testing. *J Acoust Soc Am*. 1999;106(6):L71–74. [PMID:10615713] <http://dx.doi.org/10.1121/1.428150>
36. Xin Luo, Fu QJ, Galvin JJ 3rd. Vocal emotion recognition by normal-hearing listeners and cochlear implant users. *Trends Amplif*. 2007;11(4):301–15. [PMID:18003871] <http://dx.doi.org/10.1177/1084713807305301>
37. Hawkey DJ, Amitay S, Moore DR. Early and rapid perceptual learning. *Nat Neurosci*. 2004;7(10):1055–56. [PMID:15361880] <http://dx.doi.org/10.1038/nn1315>
38. Mahncke HW, Connor BB, Appelman J, Ahsanuddin ON, Hardy JL, Wood RA, Joyce NM, Boniske T, Atkins SM, Merzenich MM. Memory enhancement in healthy older adults using a brain plasticity-based training program: a randomized, controlled study. *Proc Natl Acad Sci USA*. 2006;103(33):12523–28. [PMID:16888038] <http://dx.doi.org/10.1073/pnas.0605194103>
39. Smith GE, Housen P, Yaffe K, Ruff R, Kennison RF, Mahncke HW, Zelinski EM. A cognitive training program based on principles of brain plasticity: results from the Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study. *J Am Geriatr Soc*. 2009;57(4):594–603. [PMID:19220558] <http://dx.doi.org/10.1111/j.1532-5415.2008.02167.x>
40. Owen AM, Hampshire A, Grahn JA, Stenton R, Dajani S, Burns AS, Howard RJ, Ballard CG. Putting brain training to the test. *Nature*. 2010;465(7299):775–78. [PMID:20407435] <http://dx.doi.org/10.1038/nature09042>
41. Papp KV, Walsh SJ, Snyder PJ. Immediate and delayed effects of cognitive interventions in healthy elderly: a review of current literature and future directions. *Alzheimers Dement*. 2009;5(1):50–60. [PMID:19118809] <http://dx.doi.org/10.1016/j.jalz.2008.10.008>

Submitted for publication December 2, 2011. Accepted in revised form June 4, 2012.

This article and any supplementary material should be cited as follows:

Oba SI, Galvin JJ 3rd, Fu QJ. Minimal effects of visual memory training on auditory performance of adult cochlear implant users. *J Rehabil Res Dev*. 2013;50(1):99–110. <http://dx.doi.org/10.1682/JRRD.2011.12.0229>

ResearcherID: Sandra I. Oba, MS: A-8272-2013

