

# Benefits of Phoneme Discrimination Training in a Randomized Controlled Trial of 50- to 74-Year-Olds With Mild Hearing Loss

Melanie A. Ferguson,<sup>1</sup> Helen Henshaw,<sup>1</sup> Daniel P. A. Clark,<sup>1</sup> and David R. Moore<sup>1,2,3</sup>

**Objectives:** The aims of this study were to (i) evaluate the efficacy of phoneme discrimination training for hearing and cognitive abilities of adults aged 50 to 74 years with mild sensorineural hearing loss who were not users of hearing aids, and to (ii) determine participant compliance with a self-administered, computer-delivered, home- and game-based auditory training program.

**Design:** This study was a randomized controlled trial with repeated measures and crossover design. Participants were trained and tested over an 8- to 12-week period. One group (Immediate Training) trained during weeks 1 and 4. A second waitlist group (Delayed Training) did no training during weeks 1 and 4, but then trained during weeks 5 and 8. On-task (phoneme discrimination) and transferable outcome measures (speech perception, cognition, self-report of hearing disability) for both groups were obtained during weeks 0, 4, and 8, and for the Delayed Training group only at week 12.

**Results:** Robust phoneme discrimination learning was found for both groups, with the largest improvements in threshold shown for those with the poorest initial thresholds. Between weeks 1 and 4, the Immediate Training group showed moderate, significant improvements on self-report of hearing disability, divided attention, and working memory, specifically for conditions or situations that were more complex and therefore more challenging. Training did not result in consistent improvements in speech perception in noise. There was no evidence of any test-retest effects between weeks 1 and 4 for the Delayed Training group. Retention of benefit at 4 weeks post-training was shown for phoneme discrimination, divided attention, working memory, and self-report of hearing disability. Improved divided attention and reduced self-reported hearing difficulties were highly correlated.

**Conclusions:** It was observed that phoneme discrimination training benefits some but not all people with mild hearing loss. Evidence presented here, together with that of other studies that used different training stimuli, suggests that auditory training may facilitate cognitive skills that index executive function and the self-perception of hearing difficulty in challenging situations. The development of cognitive skills may be more important than the development of sensory skills for improving communication and speech perception in everyday life. However, improvements were modest. Outcome measures need to be appropriately challenging to be sensitive to the effects of the relatively small amount of training performed.

**Key words:** Auditory, Cognition, Learning, Speech perception, Training.

(*Ear & Hearing* 2014;35:e110–e121)

<sup>1</sup>National Institute for Health Research Nottingham Hearing Biomedical Research Unit, Nottingham, United Kingdom; <sup>2</sup>Medical Research Council Institute of Hearing Research, Nottingham, United Kingdom; and <sup>3</sup>Cincinnati Children's Hospital, Cincinnati, Ohio, USA.

Copyright © 2014 by Lippincott Williams & Wilkins. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 License, where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

## INTRODUCTION

Auditory learning, defined as improved listening through training (Moore et al. 2009), has been used since the 1950s as a clinical intervention aimed at improving communication abilities in people with hearing loss (Bamford 1981). The advent in the mid 1990s of commercial auditory training programs, such as Fast ForWord for children with language-based learning impairments (Tallal et al. 1996), provided widespread, cost-effective, easy-to-deliver training solutions that could be tailored to suit individual needs for home use. This in turn promoted a proliferation of research on individualized, computer-generated auditory training and learning. The general aim of the research was to understand the underlying principles and mechanisms of auditory training in normally hearing listeners (Wright et al. 1997; Amitay et al. 2005, 2006) and the efficacy of such interventions to improve receptive speech perception in those with hearing loss (Fu et al. 2004; Burk et al. 2006; Stecker et al. 2006). However, despite a growing increase of training products and research, we still have little clear understanding of how effective auditory training is for improving everyday listening skills.

A systematic review of the literature (Sweetow & Palmer 2005) examined the evidence that auditory training improves communication skills in adults with hearing loss. The review identified six peer-reviewed articles published between 1970 and 1996, which met the following inclusion criteria (1) randomized controlled trials (RCTs), non-RCTs, cohort, and before/after designs, with or without a control group, (2) adults with hearing loss, but not cochlear implant users, (3) training paradigm as the independent variable, and (4) outcome measures related to speech perception, or self-perception of communication abilities. It was concluded that although there was some evidence to support improved auditory skills trained during the published studies, there was no firm evidence to suggest that auditory training translated to effective, real-world benefits. The review also pointed out that these studies generally lacked scientific rigor. Four of 6 failed to include a control group, necessary to distinguish training-related improvement from test-retest effects (see also McArthur 2007), and none conducted a power calculation to define the appropriate sample size to detect clinically meaningful post-training differences.

A recent systematic review of studies since 1996 that used individual computer-based auditory training for adults with hearing loss identified 13 studies of very low to moderate study quality (Henshaw & Ferguson 2013a). Quality concerns included inadequate control for procedural learning or for placebo effects; very few of the more recent studies included a control group to assess test-retest effects. Furthermore, very few studies included a power calculation. Some did not report

results from all the outcome measures obtained, leading to a lack of transparency. Finally, blinding of participant or tester was rarely implemented. A key finding was that “on-task learning” (i.e., improvement on the trained task) usually occurred for a range of stimuli including monosyllables, syllables, words, and phrases in people with hearing loss (Burk et al. 2006; Burk & Humes 2008; Humes et al. 2009) and for both hearing aid users (Stecker et al. 2006; Sweetow & Henderson Sabes 2006; Miller et al. 2008) and cochlear implant users (Fu et al. 2004; Miller et al. 2008; Tyler et al. 2010; Oba et al. 2011).

That on-task learning occurs is interesting theoretically, and supports animal models of neuroplasticity (Recanzone et al. 1993). The evidence to support off-task or “generalization” of learning (i.e., improvements in tasks that are not trained directly) is considerably less clear (Sweetow & Palmer 2005; Henshaw & Ferguson 2013a). To date, the “gold standard” clinical test for demonstrating generalization has been improvements in speech-in-noise perception, the most common complaint of people with hearing loss. This is reflected in the majority of auditory training studies, which used speech training stimuli as well as speech outcome measures (Henshaw & Ferguson 2013a). Although there is evidence to suggest that training using multiple talkers promotes greater word-in-noise learning and that word-in-noise training can generalize to unfamiliar speakers (Burk et al. 2006), such training does not always lead to generalization to unfamiliar words, nor familiar words embedded in unfamiliar sentences (Humes et al. 2009). Training on syllables or phonemes has been shown to transfer to improvements in word-in-sentence and sentence perception in cochlear implant users (Fu et al. 2005, 2008) but not in hearing aid users (Stecker et al. 2006; Woods & Yund 2007).

While it is important to be able to demonstrate that auditory training results in measurable performance improvements, such as speech perception, it is also important for those doing the training to feel that it is benefiting them in everyday conversation, which may be best shown in self-report questionnaires. Therefore, assessment of benefit should include both subjective and objective measures (Sweetow & Henderson Sabes 2010). Our systematic review noted self-reported outcomes were used in only 3 of the 13 studies, with mixed results. Improvements were shown for hearing handicap, measured by the Hearing Handicap Inventory for the Elderly and the Communication Scale for Older Adults (Sweetow & Henderson Sabes 2010), but not by the Speech Spatial and Qualities of Hearing questionnaire (Ingvalson et al. 2013) or a health status questionnaire, the Glasgow Benefit Inventory (Stacey et al. 2010).

An increasing acknowledgment of the importance of cognition (e.g., memory and attention) in listening ability over the last 10 years has been reflected in auditory learning research. Amitay et al. (2006) showed that robust learning can occur in normally hearing adults attempting to discriminate identical tones; an impossible task. This suggests that the effects of auditory training extend beyond sensory discrimination per se, drawing upon top-down, cognitive mechanisms to improve auditory performance. This is supported by improvements in attention, auditory working memory (Stroop, listening span; Sweetow & Henderson Sabes 2006) and global auditory memory (Mahneke et al. 2006) after training on auditory stimuli.

For auditory training to be effective, those undertaking it need to comply with the intervention. As with many health-change behaviors, such as cessation of smoking and drinking alcohol

(Curry et al. 1991; DiClemente et al. 1999), compliance with behavioral interventions over relatively prolonged times can be poor, and auditory training is no exception. For example, compliance rates in the United States with the Listening and Communication Enhancement (LACE) software in a clinical population were low, at 30% (Sweetow & Henderson Sabes 2010). Historically, auditory training programs have called for prolonged training (e.g., Fast ForWord with children; typically 1 hr a day, 5 days a week for 8 weeks), but usually without any empirical evidence to support this. Besides the fact that this training is time-consuming and demotivating, it may not be necessary to train for so long. In addition, Molloy et al. (2012) have shown that for simple auditory stimuli (a frequency discrimination task) there is increased on-task learning when shorter training sessions (~8 min) rather than longer ones (>1 hr) are used. However, systematic studies of visual learning show that outcomes are, in general, related to the amount of training (Levi 2012).

One other important factor when considering auditory training as an effective clinical intervention is that speech perception performance and communication are maintained over time (Sweetow & Henderson Sabes 2006; Tyler et al. 2010; Oba et al. 2011). Current evidence suggests that post-training performance on the trained tasks does not drop back to baseline for periods of up to 6 months, and that post-training performance levels can be regained with top-up training sessions of as little as 1 hr (Burk et al. 2006).

There are still a large number of outstanding questions on the benefits of auditory training, some of which are summarized by Boothroyd (2010). These include establishing which aspects of auditory training protocols contribute to learning, how auditory training generalizes to benefits in everyday communication and quality of life, and how individual characteristics interact with training outcomes to identify candidacy for auditory training. To answer these questions with high-quality evidence, factors to be considered include the clear reporting of results (e.g., according to the CONSolidated Standards Of Reporting Trials [CONSORT] statement; see Schulz et al. 2010) and the use of outcome measures that are appropriate and sensitive (Henshaw & Ferguson 2013a). Only one study in our recent systematic review investigated the effects of auditory training on generalization to speech perception, self-report of communication difficulties and cognition (Sweetow & Henderson Sabes 2006; Henderson Sabes & Sweetow 2007). Significant improvements were seen in all three areas, although not for all individual tests. As speech perception and cognition underpin communication abilities (Kiessling et al. 2003), the main focus in the present study was to examine outcomes across speech perception, cognition, and self-report of hearing difficulties to identify whether, and how, auditory training was contributing to communication.

Most auditory training studies show highly variable outcomes across individual participants (Fu et al. 2004; Amitay et al. 2005; Humes et al. 2009; Stacey et al. 2010; Millward et al. 2011), and not everyone benefits from training (Fu et al. 2004; Oba et al. 2011). From a clinical intervention perspective, an important goal is to identify accurately who will benefit from auditory training. This could then lead to individually targeted interventions to promote effective remediation of hearing and communication difficulties, resulting in reduced disability and handicap.

Auditory training has the potential to be a useful clinical intervention to support people with hearing loss. This includes those who are hearing aid users as well as those who choose not

to wear hearing aids, or those who have mild hearing loss and would not necessarily benefit from amplification. The present phoneme discrimination training study focused on adults with mild sensorineural hearing loss who were experiencing hearing difficulties, but had not yet sought intervention for their hearing loss. The study's aims were as follows:

1. to ascertain whether phoneme discrimination training delivered improvements of trained and untrained hearing and cognitive related skills;
2. to determine whether improvements were due to learning or to test familiarity (i.e., test-retest) effects;
3. to investigate whether learning was retained after a period without training; and
4. to determine participant compliance with a home-based, computerized phoneme discrimination training program.

## PARTICIPANTS AND METHODS

This study is reported in accordance with the CONSORT statement (Schulz et al. 2010), which offers guidance for the transparent and unbiased reporting of RCTs. The CONSORT statement is intended to improve the reporting of RCTs, enabling readers to understand a trial's design, conduct, analysis, and interpretation, and to assess the validity of its results.

### Participants

Adults were initially recruited via three local Nottingham primary care practices, which sent a hearing screening questionnaire (Davis et al. 2007) to all patients who were aged 50 to 74 years (total  $n = 3326$ ) on their register. The questionnaire return rate was 42.2% ( $n = 1471$ ) of whom 1152 indicated a willingness to participate in further research. Of these, 211 people who reported hearing difficulties in both ears agreed to participate and 96 participants attended the initial test session.

A total of 44 participants (15 female, 29 male) met the inclusion criteria: (1) having symmetrical mild, sensorineural hearing loss (better ear pure-tone average thresholds between 21 and 40 dB HL across 0.5, 1, 2, and 4 kHz), (2) being a non-hearing aid user, (3) being able to run simple computer games or control a mouse if never used a computer before, and (4) having English as first language. Exclusions from the study ( $n = 52$ ) were on the basis of audiometric results ( $n = 44$ ), being an existing hearing aid user ( $n = 3$ ), unwillingness to participate ( $n = 4$ ), or inability to control a computer mouse ( $n = 1$ ).

Participants were allocated to either the Immediate Training group (IT;  $n = 23$ ) or a wait listed Delayed Training group (DT;  $n = 21$ ) by the second author using the method of minimization (Altman 1991). The grouping variables for minimization were age (younger, 50 to 62 years; older, 63 to 74 years), better hearing threshold levels (HTLs) across 0.5 to 4 kHz (better ear, 20 to 29 dB HL; poorer ear, 30 to 39 dB HL), and sex (male; female).

### Design and Study Procedure

The study used a randomized, controlled, quasi-crossover design, shown in Figure 1. Outcome measures were obtained at all visits. Test sessions are labeled so that training occurred between times  $t1$  and  $t2$ , and the retention period occurred between times  $t2$  and  $t3$  for both groups. The control (no-training) period for the DT group between  $t0$  and  $t1$  enabled assessment of test-retest effects. The auditory training software was

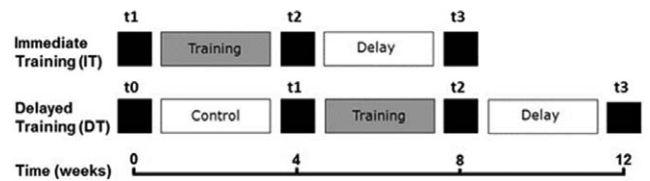


Fig. 1. Study design. Outcome measures were obtained for Immediate Training (IT) and Delayed Training (DT) groups during up to 4 visits, interspersed either with home-based phoneme discrimination training or an equivalent (control) period without training.

demonstrated to all participants in the lab at  $t1$ , before their training. The primary outcome measure was the digit triplets test. On the basis of data from the study by Wagener (2009), a power calculation to show a 2.5 dB signal to noise ratio (SNR) difference between the two groups, assuming a two-sided significance level of 0.05 and 80% power, indicated a requirement to see 20 participants in each group. On the basis of a paired-sample  $t$  test, this would result in a large effect size (Cohen's  $d = 0.89$ ).

The study was approved by the Nottingham Research Ethics Committee and Nottingham University Hospitals Trust Research and Development. Signed, informed consent was obtained. Participants were paid a nominal attendance fee and travel expenses for each visit, and a small inconvenience fee to partly recompense their time for doing the auditory training.

### Outcome Measures

**Audiological Measures** • Pure-tone air conduction thresholds (0.25, 0.5, 1, 2, 3, 4, and 8 kHz) were obtained for each ear and pure-tone bone conduction thresholds as required (0.5, 1, and 2 kHz), following the procedure recommended by the British Society of Audiology (BSA 2004), using a Siemens (Crawley, West Sussex, UK) Unity PC audiometer, Sennheiser (Hanover, Germany) HDA-200 headphones, and B71 Radioear (New Eagle, PA) transducer in a sound-attenuating booth. Otoscopy was performed and middle ear function was assessed by standard clinical tympanometry by using a GSI Tymptstar (Grason-Stadler, Eden Prairie, MN).

**Cognitive Measures** • *Nonverbal intelligence quotient* (NVIQ) was established using the Matrix Reasoning subtest of the Wechsler Abbreviated Scale of Intelligence (Wechsler 1999).

The Digit Span subtest (forward followed by backward) from the Wechsler Adult Intelligence Scale-Third Edition (Wechsler 1997) was used to measure *working memory*. Pairs of prerecorded spoken digit (0 to 9) sequences were presented at 70 dBA via Sennheiser HD-25 headphones. On successful recall of each sequence pair, the sequence increased by one digit. Discontinuation occurred when both sequences were recalled incorrectly.

The Visual Letter Monitoring task (VLM) tested *visual working memory* (Gatehouse et al. 2003). There were 10 consonant-vowel-consonant words embedded in an 80-letter sequence, and two sequences were alternated between the presentation rates across visits. Individual letters were displayed sequentially on a computer screen and participants pressed the keyboard spacebar (hit) when three consecutive letters formed a recognized consonant-vowel-consonant word (e.g., M-A-T). There were two runs, with the initial presentation rate of 1 letter/2 s, followed by 1 letter/1 s.

*Divided attention* was assessed using the Test of Everyday Attention (TEA) (Robertson et al. 1994). The Telephone Search (subtest 6; single attention) required symbols ( $n = 20$ ) to be

identified correctly, as fast as possible, while searching a simulated telephone directory. The Telephone Search While Counting (subtest 7; dual attention) required the Telephone Search to be performed while simultaneously counting strings of 1 kHz tones. The time per target for each subtest and the dual-task decrement (DTD; difference between single and dual tasks) were measured using different test versions across visits.

**Speech Perception in Noise Measures** • Two measures of speech perception in noise were presented free-field at a distance of 1 m. The *Adaptive Sentence List* (ASL) test (MacLeod & Summerfield 1990) presented sentence lists each comprising 30 items, mixed with an 8 Hz modulated noise, fixed at 60 dBA (Millward et al. 2011). Three different sentence lists were used, one at each visit. Sentences consisted of five words, including three key words (e.g., the *lunch* was *very early*), scored correct when all key words were identified. Initial sentence presentation was at 80 dBA, which varied adaptively, in 10 and 5 dB steps over two, one down-one up reversals, changing to a three down-one up paradigm using a 2.5 dB step size. The speech reception threshold (SRT) was the average SNR of the last two reversals in dB.

The *Digit Triplets test* (Smits & Houtgast 2004; Smits et al. 2005) presented series of three digits (monosyllables 0 to 9) against steady, speech-shaped background noise. Six digit lists were randomized to minimize order effects. An audibility check was performed at 65 dB SPL in quiet to ensure identification of >80%, which was increased by 5 dB until the criterion was reached. Speech level was typically 65 dB SPL. Initial digit presentation was at +5 dB SNR, and the noise varied adaptively in 2 dB steps with one down-one up reversals, and continued until 27 trials were completed. The SRT in dB was the 50% correct level.

**Self-Report Questionnaires** • The *Glasgow Hearing Aid Benefit Profile* (GHABP; Gatehouse 1999) assessed hearing disability and handicap using four predefined situations (e.g., having a conversation with 1 other person when there is no background noise; having a conversation with several people in a group) on a five-point scale (1 = no difficulty to 5 = cannot manage at all). The overall hearing Disability and Handicap scores were the mean scores converted to a percentage.

The *Speech, Spatial and Qualities of Hearing* (SSQ; Gatehouse & Noble 2004) assessed abilities and experience of hearing in different listening situations. It comprises 49 questions across three scales (1) Speech hearing (n = 14) (2) Spatial hearing (n = 17), and (3) Qualities of hearing (n = 18). Participants rated their hearing ability along a 0 to 10 visual analog scale for each question (0 = not at all to 10 = perfectly). Mean scores for each scale were derived.

**Phoneme Probe** • The discrimination threshold (%) for one phoneme continuum (/e/ and /a/) of the training task (given later in the article). Participants completed one track of 30 trials at each visit.

### Auditory Training

Home-delivered auditory training used a computer game format delivered on the IHR-STAR platform. Training was based on the “Phonomena” phoneme training package, fully described by Moore et al. (2005), but with graphics designed for adult participants. Eleven phoneme continua (/a-/uh/, /b-/d/, /d-/g/, /e-/a/, /er-/or/, /i-/e/, /l-/r/, /m-/n/, /s-/sh/,

/s-/th/, and /v-/w/), embedded in syllables where needed for natural articulation, were synthesized from endpoints consisting of real voice recordings. Each continuum transitioned from one phoneme to the other in 96 steps, saved as discrete.wav files. The stimuli were delivered through Sennheiser HD-25 headphones at a fixed level of 75 dBA. A three-interval, three-alternative, forced-choice, oddball paradigm was used; the participant’s task was to choose the odd one out from three sequentially presented phonemes. Feedback (correct/incorrect response) was given. Initially, two (identical) phonemes were selected randomly from one end of the continuum and the odd (target) phoneme from the opposite end (i.e., .wav files #1 and #96). Correct detection of the target, delivered randomly in any of the three intervals, resulted on the next trial in the identical and target phonemes being chosen from a more difficult comparison (e.g., files #11 and #86; i.e., step size 10). Trials then varied adaptively over two, one down-one up reversals, step size 10 and 5, changing to a three down-one up paradigm using a step size of 2. Performance was measured in terms of the separation between stimulus file numbers at threshold. Phoneme discrimination threshold (%) was the average of the last two reversals over 35 trials.

Phoneme pairs were selected sequentially on a rotational basis. Participants were asked to train for 15 min/day, 6 days/week over a 4-week period (360 min in total). The training was delivered, and responses logged, using a Toshiba (Weybridge, Surrey) A300 laptop, locked down to run the training program only. Two initial demonstration tasks of five trials were undertaken before home-delivered training. At the end of each training session a graphical display plotted the average threshold each day and the cumulative training time.

There was no preselection of participants based on their computer skill levels because a significant proportion of the initial sample responding to the postal questionnaire had never used a personal computer (PC; 22.1%) or the Internet (54.2%; Henshaw et al. 2012a).

### Analysis of Outcome Measures

To assess training and test-retest effects and to control for multiple testing that is implicit in repeated univariate analyses of variance, an intercept-only multivariate analysis of variance (MANOVA) for each category of key variables between the first two test sessions (IT, t1–t2; DT, t0–t1; Fig. 1) was conducted for the IT (training effect) and DT (test-retest effect) groups separately. The key variables were grouped according to speech perception (ASL sentence-in-noise; Digit Triplets), cognition (TEA, single-task decrement and DTD; VLM, 1/s and 1/2s; Digit Span) and hearing-related self-report questionnaires (GHABP, Disability and Handicap; SSQ scales, Speech and Spatial).

To assess whether the IT group demonstrated any significant training-related improvements (t1–t2) compared with the control period for the DT group (t0–t1), a between-group MANOVA of the significant measures from the previous analysis was performed, with group (IT and DT) as the between-subjects factor.

The final analysis assessed training-related effects for the whole sample. If MANOVA showed no significant difference in the pretraining and post-training results (t1–t2) between the Immediate and Delayed Training groups, the two groups were combined. A MANOVA was then performed for each set of

**TABLE 1. Mean (SD) for the on-task and off-task measures for the immediate training and delayed training groups at each visit**

Variables	Immediate Training						Delayed Training					
	Time (weeks)						Time (weeks)					
	t1 (0)	t2 (4)	t3 (8)	t0 (0)	t1 (4)	t3 (8)	t4 (12)					
On-task outcome measure												
Auditory processing												
Phoneme probe /a/-/e/ (%)	68.6 (10.5)	59.6 (4.5)	57.6 (3.5)	61.9 (6.6)	62.3 (8.9)	56.9 (7.0)	57.7 (5.9)					
Off-task outcome measures												
Speech intelligibility SRT												
Digit triplet-in-noise (dB)	-7.0 (1.3)	-7.3 (1.3)	-7.2 (1.6)	-7.0 (2.0)	-6.8 (1.6)	-7.5 (1.4)	-7.3 (1.7)					
ASL sentence-in-noise (dB)	60.5 (4.0)	59.7 (2.9)	60.1 (2.8)	62.2 (3.9)	60.8 (3.8)	60.6 (3.4)	59.0 (2.9)					
Cognition												
TEA single task	3.2 (0.8)	3.1 (0.8)	3.0 (0.7)	3.6 (0.6)	3.6 (0.8)	3.4 (0.8)	3.1 (0.6)					
TEA dual-task decrement	1.4 (1.0)	0.8 (1.0)	0.4 (0.6)	1.2 (1.0)	1.1 (1.3)	0.7 (0.9)	0.8 (0.7)					
Digit Span	16.2 (4.2)	16.8 (4.7)	16.8 (4.2)	16.6 (3.1)	17.1 (3.4)	17.7 (4.1)	18.6 (3.8)					
VLM 1 second (1/s)	5.5 (2.5)	6.7 (2.6)	6.6 (2.3)	5.8 (2.8)	6.1 (2.5)	6.7 (2.2)	7.2 (2.5)					
VLM 2 second (1/2s)	8.1 (1.7)	8.8 (1.5)	8.9 (1.5)	8.4 (1.6)	7.8 (2.2)	8.8 (2.3)	8.9 (1.2)					
Self-report												
GHABP Disability (%)	34.2 (14.9)	27.5 (13.4)	29.7 (13.5)	37.2 (17.9)	37.8 (15.1)	34.7 (17.9)	40.0 (18.6)					
GHABP Handicap (%)	27.8 (19.2)	20.8 (15.1)	21.7 (18.6)	32.3 (24.7)	31.0 (20.7)	31.6 (21.5)	31.0 (23.5)					
SSQ Speech (0-10)	5.5 (1.5)	5.8 (1.5)	5.5 (1.8)	5.3 (1.8)	5.5 (1.5)	5.4 (1.6)	5.5 (1.6)					
SSQ Spatial (0-10)	6.9 (1.5)	6.9 (1.6)	6.3 (2.0)	6.1 (1.8)	6.0 (1.8)	6.0 (1.8)	6.0 (1.8)					
SSQ Qualities (0-10)	7.3 (1.5)	7.3 (1.4)	6.8 (2.1)	6.5 (1.6)	6.7 (1.4)	6.8 (1.2)	6.8 (1.5)					

Significance levels are noted for pre- to post-training changes between visits for Immediate Training (t1–t2) and Delayed Training (t2–t3) groups. Shaded areas show the significant ( $p \leq 0.05$ ) improvements from paired *t* test analysis.

ASL, Adaptive Sentence List; GHABP, Glasgow Hearing Aid Benefit Profile; SRT, speech reception threshold; SSQ, Speech, Spatial and Qualities of Hearing; TEA, Test of Everyday Attention; VMT, Visual Letter Monitoring.

pre- to post-training outcome measures (t1–t2). Where significant training effects were shown, post hoc paired-sample *t* tests were performed to assess which individual outcome measures reached significance.

Between-group and between-visit improvements were signed to give a positive score. Effect size (Cohen's *d*) was derived for the change between visits based on the standard deviation of differences for repeated measures designs. Effect size was categorized as small, moderate, and large when Cohen's *d* was at 0.2, 0.5, and 0.8, respectively (Cohen 1988).

## RESULTS

### Performance Before Training

At baseline (IT = t1; DT = t0; Fig. 1) there were no significant demographic differences between the IT and DT groups: mean age (IT = 65.0 years; DT = 65.0 years), better ear HTL (0.5 to 4 kHz: IT = 28.2 dB HL; DT = 28.0 dB HL), socioeconomic status (Index of Multiple Deprivation; Noble et al. 2008; IT = 15,036; DT = 15,317), sex ratio (male:female: IT = 0.70; DT = 0.62), or nonverbal intelligence quotient (IT = 55.7; DT = 57.0). There was no significant difference between the groups for baseline phoneme discrimination thresholds, for any of the baseline performance tests or questionnaire scores (Table 1), or for computer skills (Table 2).

### Training Compliance

Compliance with training was high across all participants and there were no dropouts; 80% (35 of 44) of the sample completed the requested duration of training, with 75% (33 of 44) exceeding the required training. Mean training time across the

three categories of computer user (never, beginner, and competent) was 384.9, 374.3, and 379.7 min, respectively. All participants completed at least 6 full blocks, and just over two thirds (70.8%) completed at least 10 blocks. The majority of those who did not meet the requested 360 min of training were beginners ( $n = 6$ ), and the remaining three were competent PC users. All those who had never used a computer exceeded the required amount of training. There was no difference in the mean total training time for each group (IT = 377 min, SD = 50.7; DT = 378 min, SD = 46.3). The compliance rate was higher in the IT group (87%) than in the DT group (72%).

### On-Task Phoneme Learning

Across both groups of participants there was a highly significant improvement with training in phoneme discrimination threshold for all 11 phoneme pairs ( $F(1, 3931.1) = 479.1$ ,  $p < 0.001$ ), shown in Figure 2. This improvement was also evident for each group when considered separately (IT:  $F(1, 2043.3) = 153.3$ ,  $p < 0.001$ ; DT:  $F(1, 1911.2) = 84.2$ ,  $p < 0.001$ ). For each phoneme continuum, the regression line fitted to all the data points had a shallower slope than the diagonal, indicating

**TABLE 2. Computer skill mix for all participants, the Immediate Training and Delayed Training groups**

Group	N	Never	Beginner	Competent
All participants	44	7 (15.9%)	20 (45.5%)	17 (38.6%)
Immediate Training	23	3 (13 %)	11 (47.8%)	9 (39.1%)
Delayed Training	21	4 (19%)	9 (42.9%)	8 (38.1%)

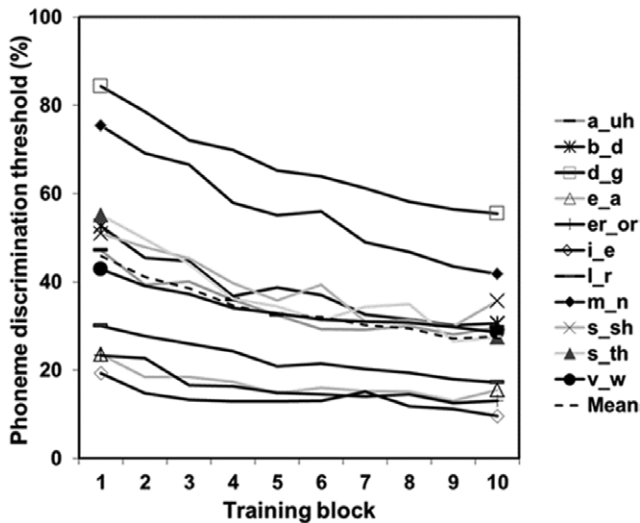


Fig. 2. Phoneme discrimination thresholds improved with training. Mean phoneme discrimination threshold values for all 11 phoneme pairs across the training period ( $n = 44$ ).

the largest improvements occurred for those individuals who had the poorest initial thresholds (Fig. 3). There was a significant correlation between the thresholds for the first and last block for

each phoneme continuum, which ranged between  $r = 0.35$  and  $r = 0.64$ , after excluding four outliers (outside the mean  $\pm 3$  SD). The vast majority of individual points fell below the diagonal, which showed that learning was evident for most participants on most of the phonemes. The overall magnitude of improvement was generally greatest for the phoneme continua that participants found most difficult to discriminate at the outset (partial  $\eta^2$ : /d/-/g/ = 0.25; /s/-/th/ = 0.24; /b/-/d/ = 0.16; /a/-/uh/ = 0.16; /m/-/n/ = 0.16; /s/-/sh/ = 0.15; /er/-/or/ = 0.14; /e/-/a/ = 0.12; /v/-/w/ = 0.12; /i/-/e/ = 0.11; /l/-/r/ = 0.10).

There was a highly significant reduction in the mean probe threshold (/e/-/a/) after training in both groups (Table 1). There was no improvement in the DT group during the no-training control phase (t0–t1), indicating that repeated testing on the probe did not itself produce improved performance (i.e., no test-retest effect). Nor did performance change for either group during the 4-week post-training period (IT: t1–t2; DT: t2–t3), indicating no further learning, or loss of learning.

### Generalization of Learning

The main analysis compared outcome measures for (1) the within-group difference between the first two visits for the IT (t1–t2) and DT groups (t0–t1) separately, and (2) the between-group difference for the first two visits (see Fig. 1).

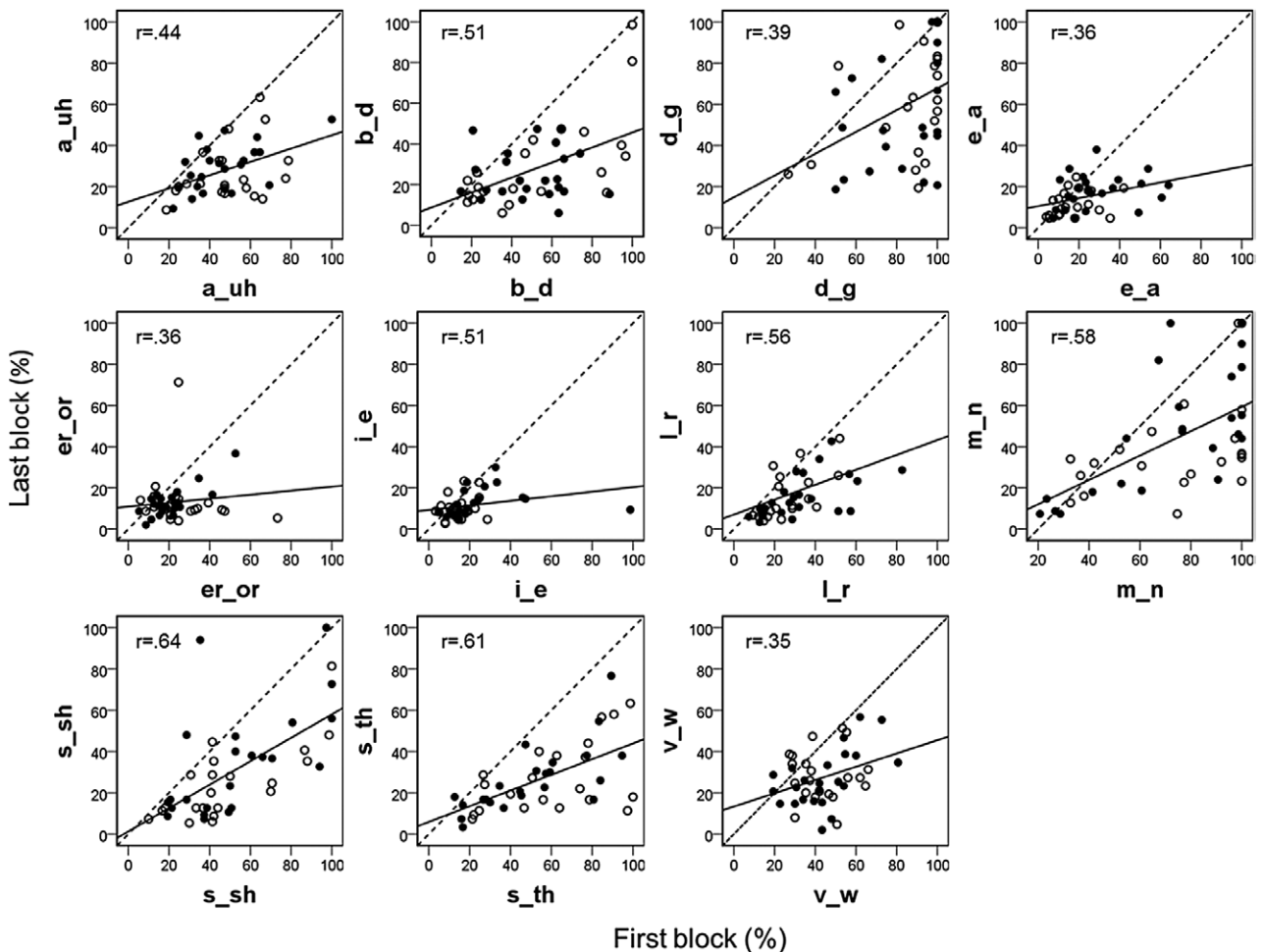


Fig. 3. The poorest initial phoneme discrimination thresholds improved the most with training. Thresholds for the first and last blocks for each individual participant. Correlation coefficient ( $r$ ) = phoneme discrimination thresholds between the first and last blocks. Solid line = regression line fitted to all the data points.

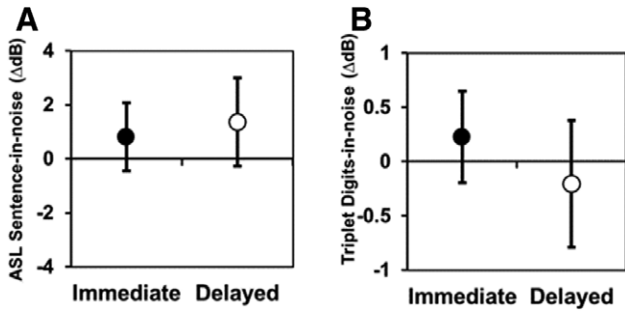


Fig. 4. Speech intelligibility did not change significantly with training or with repeated testing. Mean change ( $\Delta$ ) in SNR for (A) ASL sentence-in-noise test (B) Digit Triplets test. Data here and in Figs. 5 and 6 all show  $\Delta \pm 95\%$  confidence interval comparing performance of Immediate Training group (t1-t2) and Delayed Training group (t0-t1) (see Fig. 1).

The mean and standard deviation of the outcome measures at each test session, for both groups, are shown in Table 1.

**Speech Perception** • There was no significant between-visit change in speech-in-noise test SRTs for either the IT group (t1-t2:  $F(2, 20) = 1.02, p = 0.38$ ) or the DT group (t0-t1:  $F(2, 19) = 2.51, p = 0.11$ ), shown in Figure 4.

**Cognition** • For the IT group, MANOVA showed a significant overall improvement in performance for all the cognitive measures between t1 and t2 ( $F(5, 13) = 3.43, p = 0.03$ ). This improvement was significant for the TEA DTD ( $p = 0.02$ ), VLM for 1/s ( $VLM_{1/s}, p = 0.02$ ) and VLM for 1/2s ( $VLM_{1/2s}, p = 0.04$ ), but not for the TEA single task ( $p = 0.06$ ) or Digit Span ( $p = 0.12$ ; Fig. 5; Table 1). For the DT group, there was no change in performance between t0 and t1 ( $F(5, 10) = 0.61, p = 0.69$ ), suggesting no test-retest effects (Fig. 5; Table 1).

The between-group MANOVA across the first two test sessions for TEA DTD,  $VLM_{1/s}$  and  $VLM_{1/2s}$  showed only weak evidence to support a difference between the two groups ( $F(3, 29)$

$= 2.74, p = 0.06$ ), indicating that improvements for the IT group were not significantly greater than those for the DT group. For the combined group (IT and DT; t1-t2), there was a significant overall pre- to post-training improvement for the TEA DTD, and both VLM tasks ( $F(3, 34) = 10.35, p < 0.001$ ). Post hoc pairwise testing showed a significant effect of training for the TEA DTD ( $t(42) = 3.45, p = 0.001$ , Cohen's  $d = 0.53$ ),  $VLM_{1/s}$  ( $t(39) = 3.14, p = 0.003, d = 0.50$ ), and  $VLM_{1/2s}$  ( $t(37) = 2.10, p = 0.04, d = 0.34$ ).

**Self-Report Questionnaires** • For the IT group, MANOVA showed a significant overall within-group improvement on GHABP and SSQ scores between t1 and t2 ( $F(4, 18) = 3.25, p = 0.03$ ). This was significant for both the Disability ( $p = 0.004$ ) and Handicap ( $p = 0.031$ ) scales, shown in Figure 6A, but not for the SSQ Speech ( $p = 0.28$ ) or Spatial ( $p = 0.72$ ) scales (see Table 1). For the DT group, there was no overall within-group change between t0 and t1 ( $F(4, 17) = 0.16, p = 0.96$ ), suggesting no test-retest effect for GHABP or SSQ scales.

The between-group MANOVA across the first two test sessions for Disability and Handicap scores was not significant ( $F(2, 40) = 2.47, p = 0.09$ ). For the combined group (IT and DT; t1-t2), there was a significant pre- to post-training improvement ( $F(2, 41) = 5.87, p = 0.006$ ) for Disability and Handicap. Post hoc testing showed a highly significant effect of training for Disability ( $t(42) = 3.45, p = 0.001; d = 0.51$ ) but not for Handicap ( $t(43) = 1.53, p = 0.13; d = 0.23$ ).

Because significant benefit from training was shown for the overall Disability score, the same analysis was performed for the four individual GHABP Disability situations to assess whether improvements were dependent on situation. For the IT group, there was a significant overall within-group improvement between t1 and t2 ( $F(4, 15) = 4.0, p = 0.02$ ), which was significant only for the “having a group conversation” situation (Fig. 6B;  $p = 0.016$ ). For the DT group there was no within-group change between t0 and t1 ( $F(4, 14) = 1.34, p = 0.30$ ), suggesting no test-retest effect.

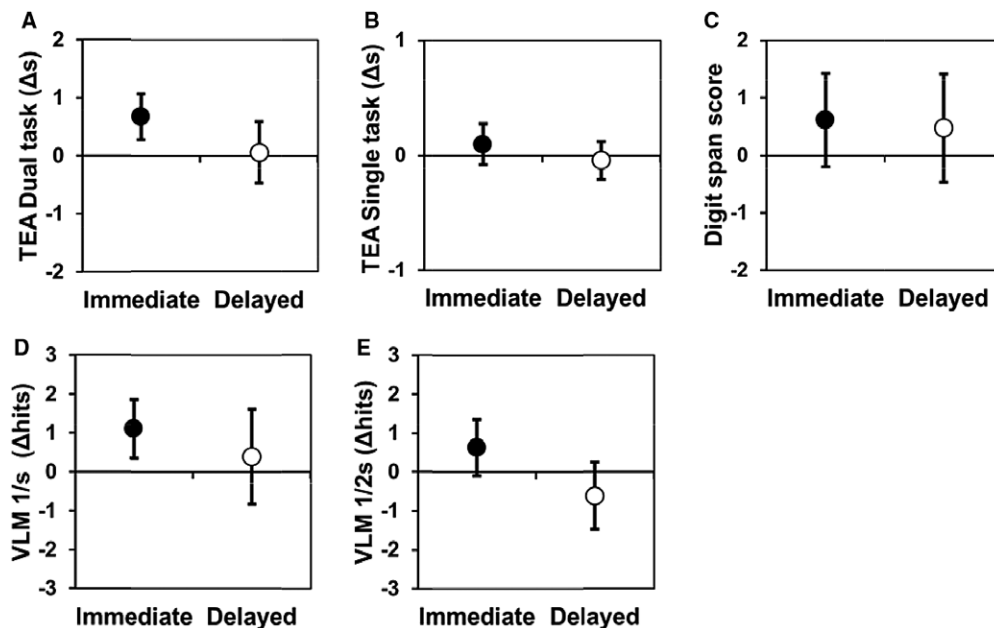


Fig. 5. Training improved complex but not simple attention and working memory. (A) Test of Everyday Attention (TEA) dual task decrement (DTD), (B) TEA single task, (C) Digit Span, (D) Visual Letter Monitoring (VLM) 1 letter/s and (E) VLM 1 letter/2s. For other details, see Fig. 4.

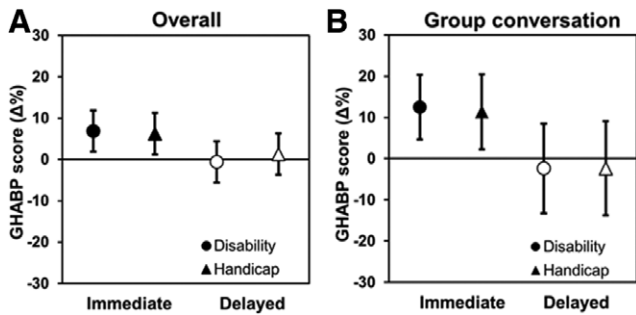


Fig. 6. Self-report of hearing disability and handicap improved with training. (A) Overall Glasgow Hearing Aid Benefit Profile (GHABP) scores, (B) GHABP "having a conversation with several people in a group." For other details, see Fig. 4.

The between-group ANOVA across the first two test sessions for the "group conversation" situation was significant ( $F(1, 42) = 4.94, p = 0.03$ ), indicating a significant improvement for the IT group compared with the DT group. For the combined group, there was a significant pre- to post-training (t1–t2) improvement for group conversation ( $t(21) = 3.17, p = 0.005; d = 0.68$ ; Fig. 6B).

**Retention of Learning**

To assess retention of learning and retention of improvements in generalizable outcome measures from training, it is assumed there is some evidence of improvement. We have defined this as any increase in performance from pre- to post-training (t1–t2). Of the measures that showed significant training-related improvement (training probe, GHABP Disability, TEA DTD, VLM tasks), between 52% and 75% of all participants showed some improvement (Fig. 7). On these tasks, significant pre- to post-training improvements were retained to t3 without further training (Fig. 7). The t3 results remained significantly better than pretraining performance (t1), and there was no significant change

during the post-training delay period (t2–t3) for any measure (see also Table 1).

We predicted that self-report would be related to performance and this was shown in the relationship between the overall GHABP Disability score with the TEA DTD. The 9 participants in the IT group who showed an improvement in both measures (Fig. 8) supported this prediction ( $r = 0.79, p < 0.01$ ). It is noteworthy that these 9 participants reported significantly greater Disability at baseline (t1) than the remaining participants in this group on overall GHABP (42.3% versus 29.0%,  $p < 0.05$ ) and SSQ Speech (4.7 versus 6.0,  $p < 0.05$ ) scores. This suggests that training-related improvements in divided attention can be predicted by poorer initial self-report on disability and speech recognition ability. There were no other factors at baseline that predicted benefit from training.

**DISCUSSION**

The overall aim of this study was to evaluate the benefits of a home-delivered, phoneme discrimination training program as a potential clinical intervention for people with mild hearing loss. A specific focus was on 50- to 74-year olds with mild sensorineural hearing loss who experienced hearing difficulties but did not have hearing aids. We found robust on-task learning of the trained phoneme continua, no improvement in speech-in-noise perception, and a mixed picture of positive and null effects on cognitive and self-report measures. Where improvements in outcome measures did occur, they were retained for at least 1 month.

Robust on-task learning was found on the trained task, consistent with many other training studies (e.g., Humes et al. 2009; Moore et al. 2009; Wright et al. 2010). In the present study, learning was apparent for all 11 phonemic contrasts, and the greatest improvement was seen for those contrasts that had the poorest performance before training at both the group mean and individual levels. Other studies have shown similar results whereby training improved the ability to discriminate difficult

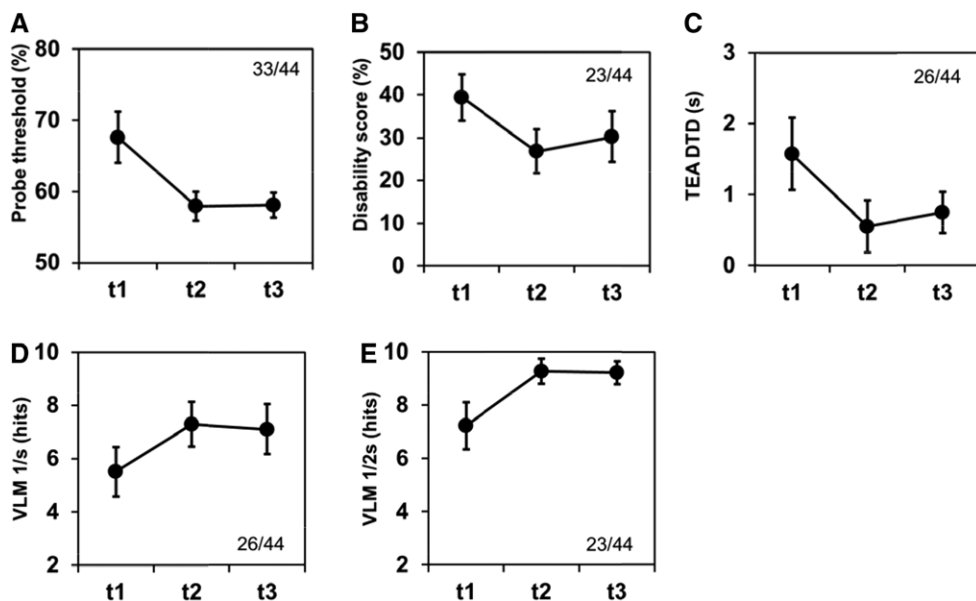


Fig. 7. Benefits of training were retained in those that showed improvements (indicated by fractions of overall participants) for both the Immediate Training and Delayed training groups for (A) Phoneme probe discrimination, (B) GHABP activity, (C) TEA dual task decrement, (D) Visual Letter Monitoring (VLM) 1/s, and (E) VLM 1/2s. Mean change (Δ) ± 95% confidence interval. t1 = pre-training, t2 = post-training, t3 = 4 weeks post-training.



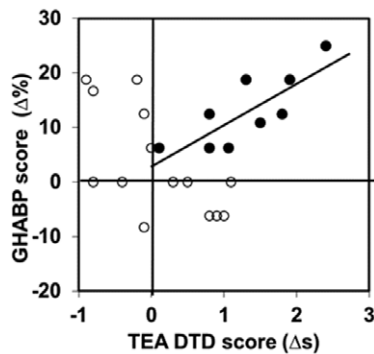


Fig. 8. Improved hearing disability and divided attention correlated following training. Filled circles: trained participants who improved on both self-rating (GHABP) of hearing disability and divided attention (TEA dual task decrement, DTD). Unfilled circles: trained participants who did not improve on at least one measure. The regression line is relevant to the filled circles only.

consonants (Stecker et al. 2006), and improvements in the perception of degraded and competing speech were greatest in those with the poorest initial scores (Henderson Sabes & Sweetow 2007). This suggests that the greatest gains on the trained task were made when initial performance was poorest.

The next and critical question for this study was whether learning transferred to improvements in untrained measures of benefit for those with mild hearing loss. As with many auditory training studies speech perception was included as a generalizable outcome but showed no significant improvement as a result of training. It may be that the high redundancy of sentence and some word stimuli reduced sensitivity to learning. The Digit Triplet test, for example, has only nine distinct speech stimuli, therefore limiting response possibilities. More generally, the evidence for transfer of learning to untrained measures of speech perception is mixed (Henshaw & Ferguson 2013a), and where it did occur there were only modest gains. Sweetow and Henderson Sabes (2006) showed the largest effect sizes occurred in the QuickSIN when presented at the more difficult presentation level of 45 dB compared with 70 dB.

This study, unlike most other training studies, examined cognition, which along with speech perception has consequences for disability and handicap arising from hearing loss. There was a consistent pattern of change in pre- to post-training performance across the cognitive measures. Significant pre- to post-training improvements, with moderate effect sizes, were seen for the complex cognitive tasks (i.e., TEA divided attention, VLM). In contrast, there were no improvements in the simple cognitive tasks (TEA single attention, Digit Span). Performance improvements were retained at 1 month post-training at similar levels to those immediately post-training, suggesting improvements were robust in the participants (approximately two-thirds) who demonstrated them. These results suggest that cognitive outcome measures need to be appropriately complex and therefore challenging to be sensitive to effects of auditory training.

Although auditory training resulted in improved performance on the complex cognitive tasks, the mechanism underlying this may not be a result of the auditory stimulus per se, but a result of active engagement with the auditory stimulus (i.e., listening). One possible explanation for the difference in observed effects for the cognitive measures used in this study is the role of executive function, an umbrella term for cognitive processes

that regulate, control, and manage other processes, such as attention, working memory, inhibition, and task-switching (Chan et al. 2008). Executive function and working memory have been shown to improve after a period of brain training (“Brain Age”) in young adults (Nouchi et al. 2013). This is consistent with our results whereby tasks that demonstrate significant post-training improvements also index executive functions (e.g., TEA divided attention [attention switching] and VLM [memory updating]). In contrast, tasks that do not demonstrate significant post-training effects do not index executive function (e.g., TEA single attention, Digit Span). The generality of this principle is further supported by evidence from a large study of multitask cognitive training in over 11,000 participants who demonstrated on-task learning but no generalizable learning on a simple Digit Span test (Owen et al. 2010). A further study to test the hypothesis that auditory training specifically improves performance on complex cognitive tasks in this population would allow a more definitive conclusion.

Of the self-report measures, training-related improvements were only demonstrated for overall hearing disability (GHABP), with a moderate effect size. Of the four individual situations that contributed to the GHABP overall score, the only significant pre- to post-training improvement, and the largest at 12.5%, was “having a conversation with several people in a group,” the most complex of the four listening situations. The other three situations improved slightly, between 2 and 6%, but none were significant. One inference from these results is that effects of training are only revealed and beneficial in listening situations that are complex, and therefore challenging. This is consistent with the cognitive results.

Effects of auditory training are often modest. In this study, pre- to post-training improvements were demonstrated within groups, but the IT group did not show significantly more improvements than the wait list DT group in the control condition. Ideally, a meta-analysis of high-quality published articles would be the best method to address the effectiveness of individual computer-based auditory training as intervention for those who have a hearing impairment. However, high variability in training stimuli (tones, syllables, words, phrases, sentences), training methods (adaptive, fixed level, user- or experimenter-controlled, home- or lab-based), outcome measures (different measures of speech perception, self-report), participant samples (hearing aid and cochlear implants users, range of hearing losses), and study quality is not currently conducive to such an approach.

Some factors that might contribute to the modest training effects include the unpredictability of task-related and procedural effects, the optimal amount of time to spend on training, and the nature of the training stimuli. In this study we demonstrated that the amount of learning varied for different training stimuli, and the proportion of participants who showed transfer of learning to generalizable outcomes varied for different outcome measures. To date, there is no clear evidence as to who would benefit from auditory training (Boothroyd 2010), although clearly this would be beneficial from a clinical perspective in terms of managing people with hearing loss. Separating procedural from perceptual learning is also problematic, with some researchers assuming that perceptual learning is a slow process requiring extensive familiarization with the training stimuli (Demany & Semal 2002; Delhommeau et al. 2002). However, others have demonstrated that perceptual learning can

be very rapid, and that procedural learning is often inaccurately confused with this early and rapid perceptual learning (Hawkey et al. 2004). Different study designs, such as the inclusion of control groups and crossover designs, can attempt to overcome or account for both task-related and procedural effects, but the uncertainty still remains. It is also unclear what the optimal duration of training is (Boothroyd 2010). It has been demonstrated that generalizable learning lags behind that of on-task learning (Wright et al. 2010). The duration of training should therefore be long enough to ensure full benefit from the transfer of learning. The vision training literature shows a clear association between the amount of training and generalizable learning effects (e.g., Levi 2012), that is, the longer the duration of training, the greater the learning. However, it has been shown for auditory frequency discrimination training that, possibly interacting with this effect, greater learning occurred with shorter rather than longer sessions (Molloy et al. 2012). The duration of training sessions in this study (15 min per day) was shorter than that used in other studies (e.g., Humes et al. 2009, 75 to 90 min per day) because it was important that the home-based training regimen would be acceptable and achievable in this group of older adults. This was confirmed in follow-up focus groups of study participants who preferred daily sessions of 15 min to alternate-day sessions of 30 min (Henshaw et al. 2012b).

A question concerning the training task was whether it was the most suitable for developing phonetic identification. This question has two aspects, whether the use of “odd-one-out” selection promotes acoustic, rather than phonological awareness, and whether training around the boundary of a categorical perception task, which these tasks were, is preferred over training within a phoneme category. It is true that a listener could potentially ignore the phonological properties of this task and perform only on the basis of discrete auditory cues. However, the rationale, especially in the present study, was that the trained listeners had auditory rather than phonological processing problems, so it was probably best to focus on a speech-based task that delivered a large number of relevant auditory discrimination trials efficiently, than on one that emphasized identification of meaningless tokens. In fact, listeners in our similar studies (Millward et al. 2011; Halliday et al. 2012) when asked about their tactics, reported discriminating whole tokens (syllables) rather than meaningless sounds. Regarding the second aspect, of training around rather than outside a categorical boundary, we reasoned that auditory discriminations would be equally difficult in both situations, and that phonetic identification would only be possible around a boundary. Note that each continuum endpoint syllable was clearly identifiable at the start of each training track.

If auditory training is to be an effective clinical intervention for people with hearing loss, it is important that the training is performed, yet participant compliance often goes unreported. Where reported, compliance rates are often exaggerated, as they are based on participant dropout rather than on those completing the required training. Only 6 of 13 studies in our recent systematic review (Henshaw & Ferguson 2013a) reported compliance figures (Stecker et al. 2006, 92.5%; Sweetow & Henderson Sabes 2006, 73%; Humes et al. 2009, 81%; Stacey et al. 2010, 73%; Oba et al. 2011, 100%; Zhang et al. 2012, 100%). These figures are comparable with the 80% found in the present study based on completion of required training, or 100% if participant dropout was set as compliance criterion. Of the other studies, only one training regimen was lab-based (Humes

et al. 2009). This suggests that those completing home-based training, where lack of supervision might be expected to result in lower compliance, are as compliant as those undergoing lab-based training. However, high compliance in volunteers who take part in a training study does not necessarily translate to high compliance in a general clinical population. Sweetow & Henderson Sabes (2010) reported compliance at 30% in a large-scale clinical trial of over 3000 participants of auditory training using LACE. It is not clear why compliance was so low but they speculate on the importance of clinician-patient interactions and patient motivation. Participants from the present study who took part in two focus groups indicated that hearing loss and the possibility of improving hearing were extrinsic motivators, whereas the desire to complete the training and to beat their previous scores during training were intrinsic motivators (Henshaw et al. 2012b). As with many health conditions, readiness to take action is required to change and improve health behaviors. The principles underpinning the Transtheoretical Health Behaviour Change Model (DiClemente & Prochaska 1998), which define a person's health behavior stage (e.g., contemplation, preparation, action, and maintenance), can also be applied to auditory training. This would form a theoretical underpinning on which further research can establish predictors to identify those who will comply with auditory training.

In the present study the most robust generalization of learning was to complex cognitive measures. Working memory is highly associated with language comprehension (Rönnberg et al. 2008). As learning is always greatest on the task that has been trained, and listening ability is also related to cognition (Moore et al. 2010; Zhang et al. 2012), these results suggest that it may be beneficial to train cognition directly. We have recently completed a trial of a working memory training program (Cogmed) in a double-blind, randomized, active-controlled trial of hearing aid users (Henshaw & Ferguson 2013b). A cognitive-based training study by Smith et al. (2009) using primarily auditory stimuli (Brain Fitness Program) showed improvements in both attention and working memory in older, though not necessarily hearing-impaired listeners, compared with an active control group (i.e., the control group had an activity to perform, in this case watching educational digital video discs). A study of working memory training has shown improvements in both memory and language (sentence repetition) skills in children with cochlear implants (Kronenberger et al. 2011). This early converging evidence suggests that to improve speech perception performance, the development of cognitive skills may be as important, or even more important than the development of sensory skills. Likewise, the development of listening or perceptual skills generally may be helped more by cognitive than by sensory training. Further research is required to inform the most effective training stimulus (auditory versus cognitive or a combination of both) to improve speech perception abilities for people with hearing loss.

## CONCLUSIONS

Significant and robust learning was demonstrated for a phoneme discrimination task in 50- to 74-year-old adults with mild hearing loss. The largest learning effects were found for the most difficult-to-discriminate phonemes. Generalization of learning was shown with moderate effect sizes for complex but not simple measures of divided attention and working memory, and

for hearing disability, specifically for complex listening conditions. There were no consistent training-related improvements in speech perception. In the participants who showed transfer of learning, the learning was retained for at least 4 weeks post-training. Compliance with home-delivered training via laptop computers in this typical, pre-hearing aid population was high, even though only one third considered they were competent PC users. In conclusion, phoneme discrimination training as used in this study provided modest self-perceived benefit for listening abilities and for complex and challenging skills that are relevant for listening in realistic environments.

### ACKNOWLEDGMENTS

The authors thank Mark Edmondson-Jones for his statistical advice, and Alison Riley and Meg Wadnerker who helped with the data collection. Special thanks to the General Practices who helped the authors recruit the participants (Leen View Surgery, Bulwell; Tolcard Hill Medical Centre, Hucknall; Wollaton Vale Health Centre, Wollaton), and finally to the participants who gave their time to help the authors.

This article presents independent research funded by the National Institute for Health Research (NIHR) Biomedical Research Unit Programme. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR or the Department of Health.

The authors declare no other conflict of interest.

Address for correspondence: Melanie Ferguson, NIHR Nottingham Hearing Biomedical Research Unit, 113 The Ropewalk, Nottingham, United Kingdom NG1 5DU. E-mail: melanie.ferguson@nottingham.ac.uk

Received December 20, 2012; accepted October 26, 2013.

### REFERENCES

- Altman, D. (1991). *Practical Statistics for Medical Research*. London, United Kingdom: Chapman & Hall.
- Amitay, S., Hawkey, D. J., Moore, D. R. (2005). Auditory frequency discrimination learning is affected by stimulus variability. *Percept Psychophys*, *67*, 691–698.
- Amitay, S., Irwin, A., Moore, D. R. (2006). Discrimination learning induced by training with identical stimuli. *Nat Neurosci*, *9*, 1446–1448.
- Bamford, J. (1981). Auditory train. What is it, what is it supposed to do, and does it do it? *Br J Audiol*, *15*, 75–78.
- Boothroyd, A. (2010). Adapting to changed hearing: The potential role of formal training. *J Am Acad Audiol*, *21*, 601–611.
- BSA. (2004). Recommended procedure for pure tone air and bone conduction threshold audiometry with and without the use of masking and determination of uncomfortable loudness levels. Retrieved from <http://www.thebsa.org.uk/docs/RecPro/PTA.pdf> [serial online].
- Burk, M. H., & Humes, L. E. (2008). Effects of long-term training on aided speech-recognition performance in noise in older adults. *J Speech Lang Hear Res*, *51*, 759–771.
- Burk, M. H., Humes, L. E., Amos, N. E., et al. (2006). Effect of training on word-recognition performance in noise for young normal-hearing and older hearing-impaired listeners. *Ear Hear*, *27*, 263–278.
- Chan, R. C., Shum, D., Touloupoulou, T., et al. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Arch Clin Neuropsychol*, *23*, 201–216.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Curry, S. J., Wagner, E. H., Grothaus, L. C. (1991). Evaluation of intrinsic and extrinsic motivation interventions with a self-help smoking cessation program. *J Consult Clin Psychol*, *59*, 318–324.
- Davis, A., Smith, P., Ferguson, M., et al. (2007). Acceptability, benefit and costs of early screening for hearing disability: A study of potential screening tests and models. *Health Technol Assess*, *11*, 1–294.
- Delhommeau, K., Micheyl, C., Jouvent, R., et al. (2002). Transfer of learning across durations and ears in auditory frequency discrimination. *Percept Psychophys*, *64*, 426–436.
- Demany, L., & Semal, C. (2002). Learning to perceive pitch differences. *J Acoust Soc Am*, *108*, 2964–2968.
- DiClemente, C. C., Bellino, L. E., Neavins, T. M. (1999). Motivation for change and alcoholism treatment. *Alcohol Res Health*, *23*, 86–92.
- DiClemente, C., & Prochaska, J. (1998). Towards a comprehensive, trans-theoretical model of change. In W. Miller & N. Heather (Eds.), *Treating Addictive Behaviours*. New York, NY: Plenum Press.
- Fu, Q., Galvin, J. J., Wang, X., et al. (2005). Moderate auditory training can improve speech performance of adult cochlear implant patients. *Acoust Res Lett Online*, *6*, 106–111.
- Fu, Q. J., & Galvin, J. J., III. (2008). Maximizing cochlear implant patients' performance with advanced speech training procedures. *Hear Res*, *242*, 198–208.
- Fu, Q. J., Galvin, J., Wang, X., et al. (2004). Effects of auditory training on adult cochlear implant patients: A preliminary report. *Cochl Impl Int*, *5* Suppl 1, 84–90.
- Gatehouse, S. (1999). Glasgow Hearing Aid Benefit Profile: Derivation and validation of client-centred outcome measures for hearing aid services. *J Am Acad Audiol*, *10*, 80–103.
- Gatehouse, S., Naylor, G., Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *Int J Audiol*, *42* Suppl 1, S77–S85.
- Gatehouse, S., & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *Int J Audiol*, *43*, 85–99.
- Halliday, L. F., Taylor, J. L., Millward, K. E., et al. (2012). Lack of generalization of auditory learning in typically developing children. *J Speech Lang Hear Res*, *55*, 168–181.
- Hawkey, D. J., Amitay, S., Moore, D. R. (2004). Early and rapid perceptual learning. *Nat Neurosci*, *7*, 1055–1056.
- Henderson Sabes, J., & Sweetow, R. W. (2007). Variables predicting outcomes on listening and communication enhancement (LACE) training. *Int J Audiol*, *46*, 374–383.
- Henshaw, H., Clark, D., Kang, S., et al. (2012a). Computer skill and internet use in adults aged 50–74 years: Influence of hearing difficulties. *J Internet Med Res*, *14*, e113.
- Henshaw, H., & Ferguson, M. A. (2013a). Efficacy of individual computer-based auditory training for people with hearing loss: A systematic review of the evidence. *PLoS One*, *8*, e62836.
- Henshaw, H., & Ferguson, M. A. (2013b). Working memory training for adult hearing aid users: study protocol for a double-blind randomized active controlled trial. *Trials*, *14*, 417.
- Henshaw, H., McCormack, A., Ferguson, M. A. (2012b). Auditory training: Exploring participant motivations, engagement and compliance. *Int J Audiol*, *51*, 263–264.
- Humes, L. E., Burk, M. H., Strauser, L. E., et al. (2009). Development and efficacy of a frequent-word auditory training protocol for older adults with impaired hearing. *Ear Hear*, *30*, 613–627.
- Ingvalson, E. M., Lee, B., Fiebig, P., et al. (2013). The effects of short-term computerized speech-in-noise training on postlingually deafened adult cochlear implant recipients. *J Speech Lang Hear Res*, *56*, 81–88.
- Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., et al. (2003). Candidature for and delivery of audiological services: Special needs of older people. *Int J Audiol*, *42* Suppl 2, S292–S301.
- Kronenberger, W. G., Pisoni, D. B., Henning, S. C., et al. (2011). Working memory training for children with cochlear implants: A pilot study. *J Speech Lang Hear Res*, *54*, 1182–1196.
- Levi, D. M. (2012). Prentice award lecture 2011: Removing the brakes on plasticity in the amblyopic brain. *Optom Vis Sci*, *89*, 827–838.
- MacLeod, A., & Summerfield, Q. (1990). A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *Br J Audiol*, *24*, 29–43.
- Mahncke, H. W., Connor, B. B., Appelman, J., et al. (2006). Memory enhancement in healthy older adults using a brain plasticity-based training program: A randomized, controlled study. *Proc Natl Acad Sci U S A*, *103*, 12523–12528.
- McArthur, G. (2007). Test-retest effects in treatment studies of reading disability: The devil is in the detail. *Dyslexia*, *13*, 240–252.
- Miller, J. D., Watson, C. S., Kistler, D. J., et al. (2008). Preliminary evaluation of the speech perception assessment and training system (SPATS) with hearing-aid and cochlear-implant users. *Proc Meet Acoust*, *2*, 1–9.
- Millward, K. E., Hall, R. L., Ferguson, M. A., et al. (2011). Training speech-in-noise perception in mainstream school children. *Int J Pediatr Otorhinolaryngol*, *75*, 1408–1417.

- Molloy, K., Moore, D. R., Sohoglu, E., et al. (2012). Less is more: Latent learning is maximized by shorter training sessions in auditory perceptual learning. *PLoS One*, *7*, e36929.
- Moore, D. R., Ferguson, M. A., Edmondson-Jones, A. M., et al. (2010). Nature of auditory processing disorder in children. *Pediatrics*, *126*, e382–e390.
- Moore, D. R., Halliday, L. F., Amitay, S. (2009). Use of auditory learning to manage listening problems in children. *Philos Trans R Soc Lond B Biol Sci*, *364*, 409–420.
- Moore, D. R., Rosenberg, J. F., Coleman, J. S. (2005). Discrimination training of phonemic contrasts enhances phonological processing in mainstream school children. *Brain Lang*, *94*, 72–85.
- Noble, M., McLennan, D., Wilkinson, K., et al. (2008). The English indices of multiple deprivation 2007. Retrieved from <http://www.communities.gov.uk/documents/communities/pdf/733520.pdf>. [serial online]. London, United Kingdom: Communities and Local Government.
- Nouchi, R., Taki, Y., Takeuchi, H., et al. (2013). Brain training game boosts executive functions, working memory and processing speed in the young adults: A randomized controlled trial. *PLoS One*, *8*, e55518.
- Oba, S. I., Fu, Q. J., Galvin, J. J., III. (2011). Digit training in noise can improve cochlear implant users' speech understanding in noise. *Ear Hear*, *32*, 573–581.
- Owen, A. M., Hampshire, A., Grahn, J. A., et al. (2010). Putting brain training to the test. *Nature*, *465*, 775–778.
- Recanzone, G. H., Schreiner, C. E., Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J Neurosci*, *13*, 87–103.
- Robertson, I. H., Ward, T., Ridgeway, V., et al. (1994). *The Test of Everyday Attention*. Bury St. Edmunds, United Kingdom: Thames Valley Test Company.
- Rönnberg, J., Rudner, M., Foo, C., et al. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *Int J Audiol*, *47* Suppl 2, S99–105.
- Schulz, K. F., Altman, D. G., Moher, D.; CONSORT Group. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMJ*, *340*, c332.
- Smith, G. E., Housen, P., Yaffe, K., et al. (2009). 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*, *57*, 594–603.
- Smits, C., & Houtgast, T. (2005). Results from the Dutch speech-in-noise screening test by telephone. *Ear Hear*, *26*, 89–95.
- Smits, C., Kapteyn, T. S., Houtgast, T. (2004). Development and validation of an automatic speech-in-noise screening test by telephone. *Int J Audiol*, *43*, 15–28.
- Stacey, P. C., Raine, C. H., O'Donoghue, G. M., et al. (2010). Effectiveness of computer-based auditory training for adult users of cochlear implants. *Int J Audiol*, *49*, 347–356.
- Stecker, G. C., Bowman, G. A., Yund, E. W., et al. (2006). Perceptual training improves syllable identification in new and experienced hearing aid users. *J Rehabil Res Dev*, *43*, 537–552.
- Sweetow, R., & Palmer, C. V. (2005). Efficacy of individual auditory training in adults: A systematic review of the evidence. *J Am Acad Audiol*, *16*, 494–504.
- Sweetow, R. W., & Henderson Sabes, J. (2006). The need for and development of an adaptive Listening and Communication Enhancement (LACE) Program. *J Am Acad Audiol*, *17*, 538–558.
- Sweetow, R. W., & Henderson Sabes, J. (2010). Auditory training and challenges associated with participation and compliance. *J Am Acad Audiol*, *21*, 586–593.
- Tallal, P., Miller, S. L., Bedi, G., et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, *271*, 81–84.
- Tyler, R. S., Witt, S. A., Dunn, C. C., et al. (2010). Initial development of a spatially separated speech-in-noise and localization training program. *J Am Acad Audiol*, *21*, 390–403.
- Wagener, K. (2009). D-1–9: Report on an optimized inventory of speech-based auditory screening & impairment tests for six languages. Available at [http://hearcom.eu/about/DisseminationandExploitation/deliverables/HearCom\\_D01-9\\_v1.pdf](http://hearcom.eu/about/DisseminationandExploitation/deliverables/HearCom_D01-9_v1.pdf) FP6-004171 HEARCOM Hearing in the Communication Society. Accessed February 14, 2013.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. New York, NY: The Psychological Corporation: Harcourt Brace & Company.
- Woods, D. L., & Yund, E. W. (2007). Perceptual training of phoneme identification for hearing loss. *Sem Hear*, *28*, 110–119.
- Wright, B. A., Buonomano, D. V., Mahncke, H. W., et al. (1997). Learning and generalization of auditory temporal-interval discrimination in humans. *J Neurosci*, *17*, 3956–3963.
- Wright, B. A., Sabin, A. T., Zhang, Y., et al. (2010). Enhancing perceptual learning by combining practice with periods of additional sensory stimulation. *J Neurosci*, *30*, 12868–12877.
- Wright, B. A., Wilson, R. M., Sabin, A. T. (2010). Generalization lags behind learning on an auditory perceptual task. *J Neurosci*, *30*, 11635–11639.
- Zhang, T., Dorman, M. F., Fu, Q. J., et al. (2012). Auditory training in patients with unilateral cochlear implant and contralateral acoustic stimulation. *Ear Hear*, *33*, e70–e79.
- Zhang, Y. X., Barry, J. G., Moore, D. R., et al. (2012). A new test of attention in listening (TAIL) predicts auditory performance. *PLoS One*, *7*, e353502.