

The Efficacy of Auditory Perceptual Training for Tinnitus: A Systematic Review

Derek J. Hoare, PhD · Paula C. Stacey, PhD ·
Deborah A. Hall, PhD

Published online: 29 July 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Auditory perceptual training affects neural plasticity and so represents a potential strategy for tinnitus management. We assessed the effects of auditory perceptual training on tinnitus perception and/or its intrusiveness via a systematic review of published literature. An electronic database search using the keywords ‘tinnitus and learning’ or ‘tinnitus and training’ was conducted, updated by a hand search. The ten studies identified were reviewed independently by two reviewers, data were extracted, study quality was assessed according to a number of specific criteria and the information was synthesised using a narrative approach. Nine out of the ten studies reported some significant change in either self-reported or psychoacoustic outcome measures after auditory training. However, all studies were quality rated as providing low or moderate levels of evidence for an effect. We identify a need for appropriately randomised and controlled studies that will generate high-quality unbiased and generalisable evidence to ascertain whether or not auditory perceptual training has a clinically relevant effect on tinnitus.

Keywords Chronic tinnitus · Perception · Handicap · Data quality

Introduction

Tinnitus aurium is a ‘phantom’ auditory experience referring to a person’s perception of a ringing, hissing or buzzing sound despite there being no such sound present in the external world. An estimated 10–15% of people in the UK population suffer from tinnitus [1] including as many as one in three older adults [2]. For a significant number of people, tinnitus has a detrimental impact on their daily lives [3]. However, tinnitus remains poorly understood and there is no uniformly effective therapy. The exploration of novel therapies for tinnitus, such as auditory training, is welcome, especially if these therapies show promise in terms of efficacy and reducing the cost of current service provision [4, 5].

Current therapies for tinnitus are based on a number of assumptions and observations. It has been proposed that tinnitus persists because the sound lacks any context or behaviourally relevant meaning that might permit the brain to ignore it over time [6, 7]. Indeed, Cuny et al. [8] suggest that without intervention, individuals fail to habituate to tinnitus and so develop a heightened awareness of it. Sound generators are one therapeutic intervention used to facilitate habituation to the tinnitus sound. These devices can be set to deliver a continuous low-level noise that helps to reduce the salience of the tinnitus and to divert attention away from it. Psychological approaches, such as relaxation therapy and cognitive behavioural therapy, are also advocated to alleviate the psychological impact of tinnitus [9]. It has been suggested that sound therapy results in little or no benefit over and above that offered by the psychological component of tinnitus therapy [10]. Tinnitus retraining therapy [11] also offers a combination of habituation and counselling, but its efficacy has not been definitively documented [12].

D. J. Hoare (✉) · D. A. Hall
National Biomedical Research Unit in Hearing, Ropewalk House,
113 The Ropewalk,
Nottingham NG1 5DU, UK
e-mail: derek.hoare@nottingham.ac.uk

P. C. Stacey · D. A. Hall
Division of Psychology, School of Social Sciences,
Nottingham Trent University,
Burton Street,
Nottingham NG1 4BU, UK

Tinnitus typically co-occurs with some degree of sensorineural hearing loss [13], suggesting that it arises from problems in the peripheral auditory system. Indeed, dominant models derived from neurophysiological data in animals consider tinnitus to be a maladaptive neuroplastic response to deafness, in terms of abnormal temporal firing [14, 15] or elevated spontaneous activity [16–18]. In cases of severe steep-sloping high-frequency hearing loss, those frequencies at the audiogram edge may become disproportionately over-represented in the brain regions that represent sound frequency, perhaps through a process of cortical rewiring or the unmasking of latent cochlear inputs to regions newly deprived of direct inputs [19, 20]. However, the supporting evidence in humans for a direct link between cortical plasticity and tinnitus is unclear at present [21].

In the audiology clinic, hearing-aid fitting is a common intervention for sensorineural hearing loss. Although the device is prescribed primarily to alleviate deafness, enhancing the sound world of patients with demonstrable hearing loss can play a very important role in alleviating tinnitus [22–25]. The neurophysiological effect of an enriched sound environment is not well established. One claim is that it prevents cortical tonotopic map reorganisation that occurs following sensory deprivation [26]. However, the research evidence from this study is not directly transferable to the audiology clinic because the research study exposed test animals to a highly unusual procedure; a continuous sound enrichment composed of multiple high-frequency tones was presented over a number of weeks, commencing immediately after the noise trauma that induced the deafness. Instead, clinical intuition favours the explanation that patients simply pay less attention to their tinnitus when external auditory stimuli are more audible [24].

It has been proposed that auditory perceptual training might provide a direct and frequency-specific method for inducing neuroplasticity by expanding the cortical representation of the trained frequency. The neural effect of repeated exposure to the same sound stimulus was first demonstrated electrophysiologically by Recanzone et al. [27]. Monkeys trained to discriminate a target frequency had significantly larger areas of primary auditory cortex tuned to that frequency than those who were untrained or those who were trained to discriminate a different target frequency. Several other animal experiments have confirmed that auditory training is associated with plastic changes in neural representations [28–31].

Given these neurophysiological perspectives, it is understandable how the ability to directly modify the cortical code for sound represents an enticing therapeutic target. The premise for perceptual training is that it facilitates the renormalisation of aberrant neural activity. Indeed, Herraiz

et al. [32] explicitly reasoned that auditory training with an active listening task would alter the cortical ‘map’ that is associated with tinnitus generation. In the somatosensory domain, a tactile discrimination training task (identifying the frequency and location of non-painful electric stimuli) was shown to be effective at reducing phantom limb pain, and this reduction in pain correlated with somatosensory cortex reorganisation as measured by electroencephalography (EEG) [33]. For tinnitus, the motivation for auditory perceptual training is that it can be tailored to redress the abnormal representations of particular frequencies in the central auditory system.

This review evaluates current evidence for a beneficial effect of auditory perceptual training on tinnitus. It addresses two objectives: (1) what is the effect of auditory training on tinnitus-related problems, and (2) could auditory training provide an effective clinical management strategy? In discussing these objectives, we also clarify the relative strengths and weaknesses of the existing body of research and suggest appropriate future research directions.

Methods

The Centre for Reviews and Dissemination (part of the National Institute for Health Research, UK) recommends core principles and methods for conducting a systematic review of health interventions [34]. This document guided our protocol for identifying the search strategy, study selection, data extraction, quality assessment and data synthesis.

Systematic Search Strategy and Study Selection

Our search strategy used electronic databases, supplemented at a later stage by searching reference lists of relevant studies and hand searching key journals. At the first stage, databases were searched using the keywords ‘tinnitus and learning’ or ‘tinnitus and training’. Databases were Cambridge Scientific Abstracts (including Medline, Biological Sciences, and Toxline), PubMed, and Web of Science. Keywords were always combined so that papers were identified only if they examined learning or training and also referred to tinnitus. Articles were selected irrespective of whether the effect of auditory perceptual training was the primary research focus. The database search in October 2009 revealed 316 articles of possible relevance. The results of this first search were screened to remove duplicates and, where possible, to also remove publications that were not available in English, were not peer-reviewed, or did not report human studies. This first evaluation procedure yielded a set of 94 articles for the next stage.

The 94 abstracts were assessed independently by two of the co-authors to select those that met a number of inclusion criteria. Inclusion criteria were identified according to the Participants, Intervention, Control, Outcome, Study design (PICOS) formula (Centre for Reviews and Dissemination [34]). In our case, *participants* were adults with chronic tinnitus, the *intervention* was an active listening task, the *controlled* design compared different types of training, a ‘trained’ group and a ‘not trained’ group, or a before and after training comparison, the *outcome* was a change in tinnitus or quality of life and the *study design* was at least a repeated measures design. Whenever an abstract contained insufficient information for evaluating it against the PICOS criteria, the article was automatically passed to the next stage for a full text review. This second evaluation procedure yielded a set of 25 articles that were judged worthy of further review. Full texts of the 25 articles were again reviewed according to the PICOS criteria for final inclusion. This third evaluation procedure excluded 16 of the articles: 4 were available only in German, 10 did not include any active listening task, 1 did not present any data and 1 did not involve participants who had tinnitus. Two of the remaining publications were from Dohrmann and colleagues [35, 36], who reported results from the same study. These two publications were treated as a single publication. This process gave eight studies for inclusion [32, 35–42]. Additional searches were informed by these selected publications and were conducted in January 2010. We searched the reference lists of our selected publications and hand-searched the content lists of relevant audiology/psychology journals published over the previous 6 months. This led to the identification of 13 further abstracts, two of which met the PICOS criteria. One was a self-citation [43] which did not include our keywords in the title or abstract, and the other was a study accessed via Epub, ahead of print [44]. We searched the reference lists of these two additional articles but did not identify any further relevant references.

Data Extraction

Data extraction was performed using a reporting form specifically developed for this purpose. Extracted data comprised the following details: (i) study design; (ii) participants (inclusion and exclusion criteria, sample size, tinnitus characteristics and hearing loss); (iii) training task and stimulus (including amount of training); (iv) compliance and follow-up (including dropout or exclusion); (v) outcome measures; and (vi) findings (outcomes, statistical comparisons, and significance levels). The reporting form for data extraction was piloted on two publications, by two co-authors, and revised as necessary. Using this revised form, data from all ten studies were then extracted

independently by two co-authors. Any differences in reporting were reconciled by jointly revisiting the relevant publication.

Quality Assessment and Data Synthesis

The aim of assessing study quality is to establish how near ‘the truth’ its findings are likely to be and whether its findings are relevant to people with tinnitus. A subset of extracted data was used to assess study quality, performed at the same time as data extraction by two co-authors. The quality assessment tool was based on Jadad et al. [45] and Oxman et al. [46]; using a procedure consistent with Ref. [34]. Eight scales provided an overall numerical quality score for each study. Seven of these eight ‘quality criteria’ represented internal validity (evidence of bias): (i) study design; validity, randomisation and control, (ii) blinding of the participants and investigators to the treatment or expected outcome, (iii) use of appropriate and validated outcome measures, (iv) matching of intervention task and control, (v) matching of participants in each group, (iv) measurement and description of compliance, and (vii) evidence of funding bias. The eighth criterion represented external validity (generalizability). For each of the eight criteria, a score of 2 was awarded if the publication met the criterion to a high standard, a score of 1 if it partially met the criterion, or a score of 0 if it was flawed or if the relevant information was not stated. Two co-authors scored studies independently and then agreed on a final score for each after discussion. This numerical scoring method gave a quality score between 0 and 16. Overall, study quality was graded as very low (0–4), low (5–8), moderate (9–12) or high (13–16). Table 1 reports the qualitative descriptors for these grades according to Oxman et al. [46].

A quantitative synthesis (meta-analysis) of comparative publications could not be carried out because the outcome measures were not sufficiently similar across all studies.

Table 1 Grading of evidence quality to support confidence in the study’s findings and estimation of effect (adapted from Oxman et al. [46])

Grading of evidence quality	Confidence and estimation of effect
High	Further studies very unlikely to change our confidence in the estimate of effect
Moderate	Further studies are likely to impact on our confidence in the estimate of effect and may change the estimate
Low	Further studies are very likely to impact on our confidence in the estimate of effect and are likely to change the estimate
Very low	The estimate of effect is very uncertain

Instead, quality assessment was incorporated into a narrative synthesis to help interpret and explain differences in results across studies.

Results

Data Extraction

A clear descriptive summary of the included studies was achieved by tabulating details of study design, participants, training task/stimulus, compliance and follow-up, outcome measures and findings (Table 2).

Study Design

The PICOS selection criteria required a repeated measures design. Hence, at minimum, studies reported before and after measurements in a tinnitus treatment group [i.e. 40, 42, 43] or an individual case study [41], without direct comparison to control participants. Five of the remaining studies report a controlled design, comparing auditory training with a different treatment group. Only one study [44] reported a randomised allocation of participants to such groups. This issue raises a key concern about internal reliability which is discussed later in the section on study quality.

Participants

Typically, both treatment and control groups comprised people with tinnitus. However, tinnitus severity was highly variable. For example, participants recruited by Herraiz et al. [44] reported a within-group tinnitus handicap ranging 32 ± 21 points, measured using the tinnitus handicap inventory questionnaire [47]. This range spans two categories of tinnitus handicap (mild and moderate). Participants who scored above 56 points (severe tinnitus) were excluded. Flor et al. [37] recruited participants presenting with symptoms which spanned three categories of tinnitus severity (slight to severe, measured using the Goebel-Hiller tinnitus questionnaire [48]). Range of tinnitus severity influences the extent to which the results are generalisable to a typical clinical caseload and the implication is discussed under the later subheading 'External validity'.

Some studies did not fully report the characteristics of the control group, and so it is not possible to determine whether the procedures for matching participants were carried out in a systematic manner. The randomisation procedure employed by Herraiz et al. [44] produced treatment and waiting list control groups that were evenly matched according to tinnitus severity, handicap and

loudness, but not according to tinnitus pitch. For example, all participants who matched their tinnitus to 3 kHz were allocated to the same group. In an earlier study [39] from the same group, participants were matched evenly according to age, gender, and hearing loss. Again, the groups were not matched according to tinnitus pitch; there were four participants in the control group whose tinnitus matched to 4 kHz, but none in the treatment group.

Training Task and Stimulus

The PICOS selection criteria required an active listening task, although the nature of the training task and stimulus differed somewhat across studies. Five studies used a single-frequency pitch discrimination task. Whether this was administered using an adaptive procedure [32, 35–37, 41] or a procedure with fixed difficulty across trials [44] appeared to have no clear influence over outcome. There has also been some investigation of whether or not the choice of target frequency determines outcome. The question regarding choice of training frequency is relevant to the goal of relating study results back to the prediction based on the neurophysiological model that perceptual training might facilitate the renormalisation of aberrant neural activity that gives rise to tinnitus. Two of the studies reported that the training frequency had no significant effect on any of the outcome measures [32, 37]. In contrast, the study by Herraiz et al. [44] concluded that training at a frequency that was similar (but not identical) to the tinnitus pitch was most effective—a result that in our opinion does not easily support the neurophysiological model. The neurophysiological model might also predict a relationship between the amount of training and the size of the benefit. Our review showed that the duration and frequency of the auditory perceptual training regime ranged across studies from 10 min twice daily for 1 month [32] to 2 h per day for 1 month [37]. However, the results indicate no reliable effect of training duration on tinnitus outcome.

Although active listening tasks were used in all ten training regimes, few authors reported measures of change in performance in these tasks. Noreña et al. [41] made repeated measures of frequency-discrimination thresholds at all four trained frequencies. They found a significant effect of training at the highest trained frequency (6.5 kHz) but not at lower training frequencies. Searchfield et al. [42] trained participants on a sound identification and localisation task whose aim was to indirectly influence tinnitus by enhancing the focus of attention towards environmental sounds. Despite recording a number of measures of attention, none showed a significant change after training.

Table 2 Descriptive synthesis of data extracted from the ten studies selected for detailed review

Reference	Study design	Participants	Training task/stimulus	Compliance and follow-up	Outcome measures	Findings
Dohrmann et al. [35, 36]	Before/after with control group (neurofeedback training)	27 (auditory training) and 21 (neurofeedback training) adults with chronic tinnitus and aided hearing loss	Frequency discrimination—adaptive difficulty, 10×2 h sessions. Pure tones in the range of hearing loss	Compliance not reported. Follow-up 6 months after training	GHTQ	No significant change in GHTQ scores after frequency discrimination training
Flor et al. [37]	Before/after with two training groups (more versus less training)	Five (more training) and seven (less training) adults with chronic tonal tinnitus and hearing loss	Frequency discrimination—adaptive difficulty, 2 h per day in <i>n</i> days over 4 weeks. Pure tones same as or different from tinnitus frequency	Non-compliance reported—due to worsening tinnitus. No reported follow-up beyond the end of training	Self-report of severity, MTI, GHTQ	After training, MTI was not significantly different. Self-report of severity and GHTQ were significantly reduced for extended training group
Herraiz et al. [38]	Before/after with control (WLG)	14 (training) and 21 (WLG) adults with tinnitus (matched between 0.5 and 8 kHz) and high-frequency hearing loss. Excluded if depressed or anxious	Auditory discrimination—fixed difficulty (4 kHz pure tones versus white noise), 10 min twice daily for 30 days	Compliance with training measured at 98%. No reported follow-up beyond the end of training	Self-report of improvement, VAS, THI	Significant improvement in severity (self-report and VAS) compared to control. THI difference was not significant
Herraiz et al. [39]	Before/after with control (WLG)	27 (training) and 26 (WLG) adults with mild to moderate tinnitus handicap. Tinnitus matched at 6 or 8 kHz. High-frequency hearing loss. Excluded if severe tinnitus handicap, depressed or anxious	Auditory discrimination—fixed difficulty (4 kHz pure tones versus white noise), 10 min twice daily for 30 days	Compliance not reported. No reported follow-up beyond the end of training	Self-report of improvement, VAS, THI	Significant improvement in severity (self-report and VAS) compared to control. Significant reduction in THI compared to control
Herraiz et al. [32]	Before/after with parallel groups training at different frequencies	11 adults per group with mild to moderate tinnitus handicap and high-frequency hearing loss	Frequency discrimination—adaptive difficulty. Pure tones either at or below tinnitus frequency, 10 min twice daily for 1 month	Compliance not reported. No reported follow-up beyond the end of training	Self-report of improvement, VAS, THI	No significant difference between training groups. Significant reduction in THI only when groups were pooled
Herraiz et al. [44]	Before/after with parallel randomised groups training at different frequencies, and a control (WLG)	20 and 21 adults per training group, 26 in WLG, mild to moderate tinnitus handicap. Excluded if severe tinnitus handicap	Frequency discrimination—fixed difficulty. Pairs of pure tones were the same or different, and at (same) or different (nonsame) to tinnitus pitch, 10 min twice daily for 30 days	Compliance fully reported. Participants who continued training after the initial month were followed up	Self-report of improvement, VAS, THI	Significant improvements in self-report and THI for trained versus WLG. Significant reduction in THI for trained (nonsame) versus trained (same).
Ince et al. [43]	Before/after without control group	Two adults, one with chronic tinnitus and high-frequency hearing loss, one with acute tinnitus and a mild notched (6 kHz) hearing loss	Adaptive loudness matching—subjects concentrated to reduce tinnitus loudness to match stimulus loudness, 9 or 12 lab sessions and home training without a stimulus	Compliance not reported fully. Followed up at 2 or 3 months after training	Tinnitus loudness (dB HTL)	Tinnitus loudness was reduced immediately after training but was not maintained
Ince et al. [40]	Before/after without control group	30 adults with tinnitus. Normal hearing to severe hearing loss	Adaptive loudness matching—subjects concentrated to reduce tinnitus loudness to match stimulus loudness, 3 to 12 lab sessions and home training without a stimulus	Compliance not reported fully. Planned a questionnaire follow-up at 6 and 12 months after training—data not available	Tinnitus loudness (dB HTL)	Significant reduction in tinnitus loudness after training

Table 2 (continued)

Reference	Study design	Participants	Training task/stimulus	Compliance and follow-up	Outcome measures	Findings
Noreña et al. [41]	Before/after of trained versus untrained ear in same participant	One adult with bilateral tonal tinnitus and high-frequency hearing loss	Frequency discrimination — adaptive difficulty. Various pure tones. 9 × 2 h sessions over 4 weeks, using one ear only	Compliance not reported fully. No reported follow-up beyond the end of training	Change in tinnitus spectrum (estimated contribution of different frequencies to tinnitus)	Difference in tinnitus spectrum (reduction in high-frequency components) in trained but not untrained ear
Searchfield et al. [42]	Before/after without control group	Ten adults with annoying tinnitus (mean = moderate tinnitus handicap) and high-frequency hearing loss	Identification and localisation of auditory objects (e.g. cat meow/left ear) in background noise. 15 days, training for 30 min per day	Compliance not reported fully. No reported follow-up beyond the end of training	Tinnitus loudness and minimum masking level	No significant change in tinnitus loudness. Significant reduction in minimum masking level

Participant numbers refer to number per group, where applicable. Only tinnitus-related outcome measures are presented

GHTQ Goebel-Hiller tinnitus questionnaire, *HTL* hearing threshold level, *MTI* multi-dimensional tinnitus inventory, *THI* tinnitus handicap inventory, *VAS* visual analogue scale, *WLG* waiting list group

Compliance and Follow-up

Compliance was under-reported. However, two of the more recent studies by Herraiz and colleagues reported a high degree of compliance with the training protocol; 98% [38] and 95% [44]. Herraiz et al. [44] also reported a low drop-out rate of 9%, whilst Flor et al. [37] reported a dropout of 14% due to worsening tinnitus. It is possible that compliance is biased by the conduct of the study. For example, neither participants nor experimenters appeared to be blind to the treatment allocation, nor was it clearly stated that participants were unaware which training group was expected to improve. The implication of any blinding bias is discussed later. Information on follow-up is also scant. There was no evidence that any of the improvements reported had any longer-term impact beyond the end of training.

Outcome Measures

Unfortunately, there are no standardised measures for evaluating the effectiveness of tinnitus interventions [49], and so it is perhaps not surprising that no two research groups have used the same set of measures. The primary outcome measure was often a self-reported rating of tinnitus severity (using a validated questionnaire or a visual analogue scale) or a personal statement regarding change in tinnitus (Table 2). Given the theoretical rationale of auditory perceptual training is the alteration of any aberrant neural activity that might be giving rise to tinnitus, then perhaps the most relevant outcome measure is a change in tinnitus percept. Only four studies used a psychoacoustic measure of tinnitus percept, namely, tinnitus loudness [40, 42, 43], pitch match [41], and minimum masking level [42].

We note that some of the studies reporting psychoacoustic measures post-training did use a number of self-reported measures at the pre-training stage [40, 42]. Additional interpretive leverage, such as levels of clinical significance, could be gained if studies assessed the efficacy of the intervention using a combination of psychoacoustic and self-reported outcome measures. A combination of positive outcomes would be the ideal goal where training is demonstrated to reduce the tinnitus sensation and this, in turn, alleviates tinnitus distress.

Findings

This section considers self-reported and psychoacoustic outcomes measures, in turn. Overall, five out of six studies reporting some type of self-reported outcome measure of tinnitus handicap or severity demonstrated a positive benefit of training. Effects of training have typically been evaluated by a repeated measures test (using the tinnitus

handicap inventory or Goebel-Hiller tinnitus questionnaire scores). Several studies have reported no reliable effect of training on Goebel-Hiller tinnitus questionnaire scores [35–37], although Flor et al. [37] did report a significant mean reduction of 5.4 Goebel-Hiller tinnitus questionnaire points for a subgroup of participants who completed the lengthiest amount of training. Herraiz and colleagues [32, 38, 39, 44] favoured the use of the tinnitus handicap inventory as an outcome measure in all four of their studies and tested the change in tinnitus handicap score across the participant groups. We discuss these results in more detail because they highlight the importance of statistical power in determining repeatability of findings. Their first study [38] found that the change in score was not significantly different for the auditory training group (mean change was -4.1 points) compared to that of the control group (mean change was $+0.1$ points). With a larger number of participants, Herraiz and colleagues [39] did show a statistically significant effect of training for the same task. However, the mean changes in tinnitus handicap inventory score differed little from those reported in the previous study (-4.9 points for the training group and $+1.5$ points for the control group) suggesting that the earlier null result might have been due to lack of statistical power. The two more recent studies by the same group [32, 44] confirmed significant effects of auditory discrimination training in terms of a 9.4- and 7.4-point reduction in tinnitus handicap inventory score, respectively.

A number of studies also evaluated self-reported outcomes using some form of non-validated personal statement of tinnitus change (e.g. ‘my tinnitus is better’ or ‘my tinnitus is worse’) [32, 37–39, 44], often in conjunction with a questionnaire-based score of tinnitus severity or handicap. However, these personal statements do not always concur with those outcomes based on the questionnaires [38]. Self-stated benefits are much more likely to demonstrate an effect than a comparable validated outcome measure. For example, Marshall [50] reported a 40% bias of positive results in randomised controlled trials where self-styled outcome measures were used instead of validated measures.

All four studies reporting some type of psychoacoustic outcome measure demonstrated a significant change. Changes in tinnitus loudness, matched frequency spectrum, and minimum masking level have again typically been evaluated by a repeated measures test. Noreña et al. [41] reported a significant change of tinnitus percept in the trained ear after frequency-discrimination training. After training, a reduction in the contribution of high-frequency components to the participant’s tinnitus was observed, whereby frequencies above 8 kHz no longer contributed to the reported spectrum. This effect might be interpreted as a benefit because it in effect, completely extinguished a

component of the participant’s tinnitus. The authors did, however, speculate that training might have refined the participant’s matching of their tinnitus (i.e. changed cognitive representations) rather than changing the tinnitus itself.

One study evaluated training using more than one psychoacoustic outcome measure and so it is informative to explore the degree of convergence between results. Although Searchfield et al. [42] reported a significant reduction in the average minimum level of a noise required to mask tinnitus of 13.2 dB, training did not show a concomitant reduction in tinnitus loudness. This discrepancy is unexpected and perhaps indicates that at least one of the outcome measures is not reliable. Certainly, participants showed a large variability in terms of the effect of training on tinnitus loudness. The findings reported by Ince et al. [40, 43] also highlight some temporal instability in the measure, since there was a large within-session reduction in tinnitus loudness, which did not always carry over to the subsequent training session.

Study Quality and Data Synthesis

Numerical quality scores allocated for the eight quality criteria are summarised in Table 3 and are discussed item by item below.

Study Design

Randomisation is the preferred or ‘gold standard’ design in therapeutic or ‘effectiveness of an intervention’ research [46, 51, 52]. Only one study, from Herraiz et al. [44], was reported as a randomised controlled trial, where participants were randomly allocated to different treatment groups (scoring a 2 for this criterion). Without randomisation, one cannot be confident that the observed effects are truly attributable to the auditory training intervention, rather than to intrinsic differences between groups or spontaneous recovery over time. Other studies scored 1 because they met the minimum criterion of a repeated measures design. Although studies implemented a repeated measures design, most of the authors did not confirm that the measure was stable before the start of treatment. Without any evidence of the test–retest reliability, we cannot be certain that any changes observed during training are specific to the intervention. Alternative interpretations of change might be an ‘anticipation of treatment’ effect or a ‘therapeutic relationship’ effect and these cannot be ruled out.

Blinding

Blinding of the participants and the experimenters to the treatment allocation is the most preferable or ‘gold

Table 3 Assessment of eight quality criteria, use of power calculations, type of outcome measures, and the resulting overall quality rating for each of the ten studies

Reference	Study design	Blinding	Outcome measure	Matched training task and control	Matched participant groups	Compliance reported	Evidence of funding bias	Externally valid	Use power calculation	Outcome measures	Quality of evidence
Dohrmann et al. [35, 36]	1	0	1	2	1	0	2	1	X	S	Low
Flor et al. [37]	1	1	1	2	0	1	2	1	X	S	Moderate
Herraiz et al. [38]	1	0	1	1	0	1	2	1	X	S	Low
Herraiz et al. [39]	1	0	1	1	2	0	2	1	X	S	Low
Herraiz et al. [32]	1	0	1	2	1	0	2	1	X	S	Low
Herraiz et al. [44]	2	0	1	2	2	2	2	1	X	S	Moderate
Ince et al. [43]	1	0	1	0	0	1	2	0	X	P	Low
Ince et al. [40]	1	0	1	0	0	1	2	1	X	P	Low
Noreña et al. [41]	1	0	1	1	1	1	2	0	X	P	Low
Searchfield et al. [42]	1	0	1	0	0	0	2	1	X	P	Low

2 Study meets criterion to a high standard, 1 study partially meets criterion, 0 flawed or the relevant information was not stated, X no or not stated, S self-reported outcome measure, P psychoacoustic outcome measure

standard' design that would achieve a quality score of 2. No study reported the proper use of participant and experimenter blinding, and most scored zero on this criterion. One study described a partial degree of blinding, but this was limited to their neurofeedback group and did not apply to their auditory training group [35, 36]. We allocated a score of 1 to the study by Flor et al. [37]. In this case, it is reasonable to assume that the participants and experimenters were unaware of the group classification since this particular outcome was the result of a post hoc reallocation of participants according to the amount of training completed.

Outcome Measures

Since none of the studies took the opportunity to directly associate physiological change (as evidenced by a change in tinnitus sensation) with a change in a measure of tinnitus handicap or severity, all scored 1 for this criterion.

Matched Training Task and Control

Most studies report some form of control for the training task, either a group who did not train (scored as 1), a group who completed a different training regime (scored as 2), or in the case of Noreña et al. [41], the untrained ear of the trained participant (scored as 1). Some concern about the appropriateness of the control has been expressed by a number of the authors. In those studies where all participant groups complete some form of potentially therapeutic intervention, there is little control over those benefits that might be due to non-specific (e.g. interpersonal) factors or placebo. Flor et al. [37] acknowledge this limitation. In studies from Herraiz et al. [38, 39, 44] where the control group remained on the waiting list for the duration of the 'treatment' period, there is again no control over non-specific factors such as spontaneous recovery. In their two most recent publications [32, 44], Herraiz and colleagues expressed the future intention to exert better experimental control by requiring the control group to complete a passive listening task or an unrelated active listening task. This control is to be welcomed as it will strengthen the quality of evidence for a *specific* auditory training benefit.

Matched Participant Groups

Studies not reporting a control group scored 0 for this criterion. Studies that reported control groups but did not provide sufficient detail to allow comparison with the intervention group were scored as 1 [32, 35, 36]. Noreña et al. [41] was also scored as 1 as the trained and untrained condition referred to the same individual. Only two studies [37, 39] provided sufficient detail to measure similarity

between groups and hence scored 2. In terms of internal validity, it is interesting to note here the general inconsistency in quality, even between studies conducted by the same author. One example relates to the quality of reporting control group characteristics by Herraiz and colleagues in three studies [32, 38, 39] which scored 1, 0 and 2, respectively.

Compliance Reported

Only Herraiz et al. [44] provided a comprehensive description of compliance, scoring 2 for this criterion. Five studies reported on some aspect of compliance, although not always in a quantifiable way, scoring 1. Compliance was not reported in four studies, scoring 0.

Evidence of Funding Bias

None of the reviewed studies suggested any evidence of a potential bias relative to their funding source, all therefore scored 2 for this criterion.

External Validity

Case reports scored 0 for this criterion, whilst most group studies scored 1 because they involved a clinical cohort. Studies recruiting sufficient numbers of participants with mild to moderate tinnitus have greater external validity than those recruiting people with severe tinnitus, because 66% of a typical clinical caseload reflects mild to moderate tinnitus handicaps (i.e. tinnitus handicap inventory scores between 18 and 56), whilst only 17% reflects severe tinnitus handicaps (tinnitus handicap inventory scores >56) [53].

Despite appropriate participant characteristics, none of the studies reported the use of a power calculation to determine sample size. Small sample sizes limit how well findings can be generalised to a typical caseload because it is not possible to accurately estimate the within- and between-group variance. Two example studies estimate the required sample size for reliably assessing tinnitus interventions. Landgrebe et al. [54] determined that a sample size of 68 participants per group was required to show a statistical and clinically relevant change in Goebel-Hiller tinnitus questionnaire score, more than twice the number studied by Dohrmann et al. [35, 36]. Using tinnitus handicap inventory scores to calculate their required sample, Gudex et al. [53] estimated that they needed 50 participants per group to be confident of a statistically significant and clinically relevant change in tinnitus handicap. None of our selected studies using the tinnitus handicap inventory as an outcome measure recruited treatment groups of this size. For Newman and colleagues [55], a change of at least 20 points in an individual tinnitus

handicap inventory score was needed to be considered ‘clinically relevant’ because the 95% confidence interval is of that magnitude. Again, none of our included studies using the tinnitus handicap inventory as an outcome measure reported score changes of this magnitude.

Study quality was evaluated by comparing the overall score with the grading presented in Table 1. The overall scores enable us to address the first objective concerning the evidence that auditory training has an effect on tinnitus-specific outcome measures. The studies were judged to provide either low or moderate levels of evidence that auditory perceptual training may be of benefit in relieving tinnitus symptoms. Hence, we must conclude that further studies of auditory training interventions are likely to impact on our confidence in the estimate of the treatment effect and may even change the estimate of that effect [46].

Discussion

The first objective of this systematic review was to evaluate the evidence that auditory training influences tinnitus-related problems. Overall, 9 out of the 10 studies reported a statistically significant change. However, our preceding narrative synthesis of the quality assessment concluded that this evidence is not of sufficient quality to be confident that auditory training represents a strategy for improving the percept and severity or handicap of tinnitus. The evidence is simply not yet robust enough to guide treatment.

Turning now to the second objective concerning whether auditory training could provide an effective clinical management strategy, we discuss key issues that are fundamental to determining the efficacy of auditory perceptual training for tinnitus and, subsequently, to realising its development into a clinically useful tool. Whilst most studies reported some significant effect of auditory training on their chosen outcome measures, none explored the implications for their finding on the desired endpoint of this line of research, i.e. tangible benefit for the person who has tinnitus. This leaves open questions about the nature of any translational benefit of auditory training for tinnitus, and whether any benefits are specific to *auditory* training. We suggest a number of practical recommendations for future research to be considered when designing a high-quality randomised controlled trial.

First, it is important to evaluate whether or not the appropriate outcome measure is being used. The selection of a primary outcome measure should be driven by the desired observation. McFerran and Baguley [56] propose that, from the perspective of the audiologist, inhibiting tinnitus should be the aim of clinical intervention. Such a viewpoint would suggest a psychoacoustic measure of tinnitus (such as tinnitus loudness or minimum masking

level) as the primary outcome measure rather than a self-reported measure (such as a reduction in tinnitus distress). However, based on this review, we recommend that a combination of psychoacoustic and self-reported outcome measures is implemented so that we can understand how the intervention reduces tinnitus percept and alleviates tinnitus distress.

Second, confidence in the stability of pre-training measures is crucial in order to be able to attribute patient benefit to the specific intervention. Perhaps this issue is of utmost concern for self-reported outcome measures using questionnaires about tinnitus handicap or distress, because clinical intuition suggests that the psychological response often naturally diminishes over time. For example, Dohrmann et al. [35, 36] provide information on the test–retest reliability of the Goebel-Hiller tinnitus questionnaire by measuring tinnitus severity twice before training. The score reduced from 34.5 to 29.5, with no further change after training. A drop of 5 points equates to a clinically meaningful shift in tinnitus severity from moderate to mild, which the authors attribute to an ‘anticipation of treatment’ effect. Hence, designs that do not confirm a stable baseline measure cannot attribute any change after auditory training to be a specific benefit of the intervention. To rule out non-specific benefits, we recommend applying a test–retest procedure during the pre-training stage or using a crossover comparison design.

Third, it is important to evaluate whether or not the outcome reflects a clinically significant change that shifts someone from one category of tinnitus severity to a less severe category. A change in Goebel-Hiller tinnitus questionnaire score of 5 points is considered clinically relevant and observable in a clinical population [54, 57]. Indeed, the two studies using this questionnaire reported a mean change of at least this magnitude [35–37]. Herraiz et al. [32, 38, 39, 44] favoured the tinnitus handicap inventory, and a change of between 4.9 and 9.4 was reported as significant between the training and control groups. The tinnitus handicap inventory is a well-validated questionnaire for an initial assessment of tinnitus severity, but is less sensitive as a measurement tool for treatment outcome. The 95% confidence interval of the tinnitus handicap inventory score is 20 points [55], meaning that a change of at least 20 points is required for the difference to be considered clinically significant. Hence, across the studies from Herraiz et al., auditory perceptual training would appear not to facilitate a significant *clinical* benefit. Moreover, details of individual benefit were not provided, so any indication of how many participants reduced their tinnitus handicap by 20 or more cannot be assessed. In the future, we hope that the work from Henry et al. [58] will deliver a standardised tinnitus outcome measure for clinicians and researchers. These authors are developing

a tinnitus functional index questionnaire, specifically to have high discriminative and evaluative validity for the assessment of tinnitus treatment outcome. Until this becomes available, we advise against the use of the tinnitus handicap inventory as a self-reported outcome measure of *change*.

Fourth, given the commitment of time and resources that auditory perceptual training can involve, it is important that any benefit is proven to be maintained beyond the training period. Longer-term outcomes are largely unreported in the reviewed studies and follow-up assessment should be a priority for future research.

Fifth, research to date has focused on the alleviation of tinnitus symptoms and has ignored usability issues relating to the training technologies. Whilst proof of concept is certainly an essential part of this aspect of translational research, to be broadly successful, any training regime must have intrinsic motivation [59, 60]. Indeed, a recent review of the efficacy of auditory training in adults highlighted that training needs to be engaging and to promote a desire to train through feedback or reward [61].

Conclusion

To date, the published evidence describing the effects of auditory training interventions for tinnitus is of low to moderate quality. We have therefore identified a need for randomised controlled studies that will generate high-quality unbiased and generalisable evidence for whether or not auditory perceptual training has a clinically relevant effect on tinnitus. Only when this evidence exists can particular forms or regimes of training be conclusively included or excluded from future research, and there can be a move towards large-scale clinical trials. As it stands, the question of whether auditory training might be developed as a clinical tool to manage tinnitus remains open to future findings. Our perspective on future research makes the following recommendations:

- Use a combination of psychoacoustic and self-reported outcome measures
- Establish confidence in the stability of the pre-training measures
- Use a self-reported outcome measure that is sensitive to change (i.e. not the tinnitus handicap inventory)
- Demonstrate longer-term benefits
- Ensure the training regime is intrinsically motivating

Acknowledgement The National Biomedical Research Unit in Hearing (DJH, DAH) is funded by the National Institute for Health Research.

Conflict of Interest None

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Davis A, El Rafaie A, editors. Tyler's tinnitus handbook. San Diego: Singular Publishing Group; 2000
- Sataloff J, Sataloff RT, Lueneburg W. Tinnitus and vertigo in healthy senior citizens without a history of noise exposure. *Am J Otol.* 1987;8:87-9.
- Bartels H, Middel BL, van der Laan BFAM, Staal MJ, Albers FWJ. The additive effect of co-occurring anxiety and depression on health status, quality of life and coping strategies in help-seeking tinnitus sufferers. *Ear Hear.* 2008;29:947-56.
- Aazh H, Moore BCJ, Roberts P. Patient-centred tinnitus management tool: A clinical audit. *Am J Audiol.* 2009;18:7-13.
- Newman CW, Sandridge SA, Meit SS, Cherian N. Strategies for managing patients with tinnitus: A clinical pathway model. *Semin Hear.* 2008;29:300-9.
- Jastreboff PJ, Gray WC, Gold SL. Neurophysiological approach to tinnitus patients. *Am J Otol.* 1996;17:236-40.
- Searchfield GD, Munoz DJB, Thorne PR. Ensemble spontaneous activity in the guinea-pig cochlear nerve. *Hear Res.* 2004;192:23-35.
- Cuny C, Noreña A, El Massioui F, Chery-Croze S. Reduced attention shift in response to auditory changes in subjects with tinnitus. *Audiol Neurotol.* 2004;9:294-302.
- Andersson G. The role of psychology in managing tinnitus: A cognitive behavioural approach. *Semin Hear.* 2001;22:65-76.
- Mckenna L, Irwin R. Sound therapy for tinnitus—sacred cow or idol worship?: An investigation of the evidence. *Audiol Med.* 2008;6(1):16-24.
- Jastreboff PJ, Jastreboff MM. Tinnitus retraining therapy (TRT) as a method for treatment of tinnitus and hyperacusis patients. *J Am Acad Audiol.* 2000;11:156-61.
- Phillips JS, McFerran D. Tinnitus retraining therapy (TRT) for tinnitus. Cochrane database of systematic reviews. 2010; in press
- Hoffman HJ, Reed GW, eds. *Snow's tinnitus: Theory and management*. Hamilton: BC Decker; 2004.
- Ochi K, Eggermont JJ. Effects of quinine on neural activity in cat primary auditory cortex. *Hear Res.* 1997;105:105-18.
- Noreña AJ, Eggermont JJ. Changes in spontaneous neural activity immediately after an acoustic trauma: Implications for neural correlates of tinnitus. *Hear Res.* 2003;183:137-53.
- Eggermont JJ, Komiya H. Moderate noise induced trauma in juvenile cats results in profound cortical topographic map changes in adulthood. *Hear Res.* 2000;142:89-101.
- Kimura M, Eggermont JJ. Effects of acute pure tone induced hearing loss on response properties in three auditory cortical fields in cat. *Hear Res.* 1999;135:146-62.
- Bauer CA. Mechanisms of tinnitus generation. *Curr Opin Otolaryngol Head Neck Surg.* 2004;12:413-7.
- Rajan R, Irvine DRF, Wise LZ, Heil P. Effects of unilateral cochlear lesions in adult cats on the representation of lesioned and unlesioned cochleas in primary auditory cortex. *J Comp Neurol.* 1993;338:17-49.
- Robertson D, Irvine DRF. Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J Comp Neurol.* 1989;282:456-71.
- Adjamian P, Sereda M, Hall DA. The mechanisms of tinnitus: Perspectives from human functional neuroimaging. *Hear Res.* 2009;253:15-31.
- Kochkin S, Tyler R. Tinnitus treatments and the effectiveness of hearing aids: Hearing care professional perceptions. *Hear Rev.* 2008;15:14-8.
- Pulec JL, Hodell SF, Anthony PF. Tinnitus: Diagnosis and treatment. *Ann Otol Rhinol Laryngol.* 1978;87:821-33.
- Trotter MI, Donaldson I. Hearing aids and tinnitus therapy: A 25-year experience. *J Laryngol Otol.* 2008;122:1052-6.
- Willott JF. Physiological plasticity in the auditory system and its possible relevance to hearing aid use, deprivation effects, and acclimatization. *Ear Hear.* 1996;17:66S-77S.
- Noreña AJ, Eggermont JJ. Enriched acoustic environment after noise trauma reduces hearing loss and prevents cortical map reorganization. *J Neurosci.* 2005;25:699-705.
- Recanzone GH, Schreiner CE, Merzenich MM. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J Neurosci.* 1993;13:87-103.
- Ohl F, Scheich H. Differential frequency conditioning enhances spectral contrast sensitivity of units in auditory cortex (field AI) of the alert Mongolian gerbil. *Eur J Neurosci.* 1996;8:1001-17.
- Kilgard MP, Merzenich MM. Cortical map reorganization enabled by nucleus basalis activity. *Science.* 1998;279:1714-8.
- Gao E, Suga N. Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system. *Proc Natl Acad Sci U S A.* 1998;95:12663-70.
- Syka J. Plastic changes in the central auditory system after hearing loss, restoration of function, and during learning. *Physiol Rev.* 2002;82:601-36.
- Herraiz C, Diges I, Cobo P, Aparicio JM. Cortical reorganisation and tinnitus: principles of auditory discrimination training for tinnitus management. *Eur Arch Otorhinolaryngol.* 2009;266:9-16.
- Flor H, Denke C, Schaefer M, Grusser S. Effect on sensory discrimination training on cortical reorganisation and phantom limb pain. *Lancet.* 2001;357:1763-4.
- Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. University of York, York. 2009. http://www.york.ac.uk/inst/crd/systematic_reviews_book.htm. Accessed 23 April 2010
- Dohrmann K, Elbert T, Schlee W, Weisz N. Tuning the tinnitus percept by modification of synchronous brain activity. *Restor Neurol Neurosci.* 2007;25:371-8.
- Dohrmann K, Weisz N, Schlee W, Hartmann T, Elbert T. Neurofeedback for treating tinnitus. *Prog Brain Res.* 2007;166:473-86.
- Flor H, Hoffmann D, Struve M, Diesch E. Auditory discrimination for the treatment of tinnitus. *Appl Psychophysiol Biofeedback.* 2004;29:113-20.
- Herraiz C, Diges I, Cobo P, Plaza G, Aparicio JM. Auditory discrimination therapy (ADT) for tinnitus management: Preliminary results. *Acta Otolaryngol.* 2006;126:80-3.
- Herraiz C, Diges I, Cobo P. Auditory discrimination therapy (ADT) for tinnitus management. *Prog Brain Res.* 2007;166:467-71.
- Ince LP, Greene RY, Alba A, Zaretsky HH. A matching-to-sample feedback technique for training self-control of tinnitus. *Health Psychol.* 1987;6:173-82.
- Noreña A, Micheyl C, Chéry-Croze, Collet L. Psychoacoustic characteristics of the tinnitus spectrum: Implications for the underlying mechanisms of tinnitus. *Audiol Neurotol.* 2002;7:358-69.
- Searchfield GD, Morrison-Low J, Wise K. Object identification and attention training for treating tinnitus. *Prog Brain Res.* 2007;166:441-60.

43. Ince LP, Greene RY, Alba A, Zaretsky HH. Learned self-control of tinnitus through a matching-to-sample feedback technique: A clinical investigation. *J Behav Med.* 1984;7:355-65.
44. Herraiz C, Diges I, Cobo P, Aparicio JM, Toledano A. Auditory discrimination training for tinnitus treatment: the effect of different paradigms. *Eur Arch Otorhinolaryngol.* 2010; doi:10.1007/s00405-009-1182-6
45. Jadad AR, Moore A, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clin Trials.* 1994;17:1-12.
46. Oxman AD, GRADE working group. Grading quality of evidence and strength of recommendations. *Br Med J.* 2004;328:1490-4.
47. Newman CW, Jacobson GP, Spitzer JB. Development of the tinnitus handicap inventory. *Arch Otolaryngol Head Neck Surg.* 1996;122:143-8.
48. Goebel G, Hiller W. The tinnitus questionnaire. A standard instrument for grading the degree of tinnitus. Results of a multicenter study with the tinnitus questionnaire (TF). *HNO.* 1994; 42:166-72
49. Langguth B, Goodey R, Azevedo A, et al. Consensus for tinnitus patient assessment and treatment outcome measurement: tinnitus research initiative meeting, Regensburg, July 2006. *Prog Brain Res.* 2007;166:525-36.
50. Marshall M. Unpublished rating scales: A major source of bias in randomised controlled trials of treatments for schizophrenia. *Br J Psychiatry.* 2000;176:249-52.
51. Dobie RA. A review of randomized clinical trials in tinnitus. *Laryngoscope.* 1999;109:1202-11.
52. Moher D, Sampson M, Campbell K, et al. Assessing the quality of reports of randomized trials in paediatric complementary and alternative medicine. *BMC Pediatrics.* 2002;2:2. doi:10.1186/1471-2431-2-2.
53. Gudex C, Skellgaard PH, West T, Sørensen J. Effectiveness of a tinnitus management programme: A 2-year follow-up study. *BMC Ear Nose Throat Disord.* 2009. doi:10.1186/1474-6815-9-6.
54. Landgrebe M, Binder H, Koller M, et al. Design of a placebo-controlled, randomized study of the efficacy of repetitive transcranial magnetic stimulation for the treatment of chronic tinnitus. *BMC Psychiatry.* 2008;8:23.
55. Newman CW, Sandridge SA, Jacobson GP. Psychometric adequacy of the tinnitus handicap inventory (THI) for evaluating treatment outcome. *J Am Acad Audiol.* 2001;12:379-80.
56. McFerran DJ, Baguley DM. Is psychology really the best treatment for tinnitus? *Clin Otolaryngol.* 2009;34:99-101.
57. Kleijnung T, Steffens T, Sand P, et al. Which tinnitus patients benefit from transcranial magnetic stimulation? *Otolaryngol Head Neck Surg.* 2007;137:589-95.
58. Henry JA, Meikle M, Griest S, Stewart B. Development of a new questionnaire: the tinnitus functional index. Proceedings of the IXth international tinnitus seminars, Department of Audiology, Sahlgrenska Academy, Göteborg, Sweden. 2008
59. Malone TW, Lepper M, editors. Snow RE, Farr MJ, Farr MJ. Aptitude, learning and instruction. Vol. 3: Cognitive and affective process analyses. London: Routledge; 1987
60. Kenny RF, Gunter GA. Endogenous fantasy-based serious games: Intrinsic motivation and learning. *Int J Soc Sci.* 2007;2:8-13.
61. Sweetow R, Palmer CV. Efficacy of individual auditory training in adults: A systematic review of the evidence. *J Am Acad Audiol.* 2005;16:494-504.