A 'Music-Related Quality of Life' (MuRQoL) measure to guide music rehabilitation for adult CI users

Giorgos Dritsakis[1], Rachel M. van Besouw[1], Pádraig Kitterick[2] & Carl A. Verschuur[3]

[1] ISVR, University of Southampton, [2] National Institute for Health Research, Nottingham Hearing Biomedical Research Unit, Nottingham, UK, [3] University of Southampton Auditory Implant Service

Abstract

Purpose: A music-related quality of life (MuRQoL) questionnaire was developed for the evaluation of music rehabilitation for adult cochlear implant (CI) users. The present studies were aimed at refinement and validation. **Method:** Twenty-four experts reviewed the MuRQoL items for face validity. A refined version was completed by 147 adult CI users and psychometric techniques were used for item selection, assessment of reliability and definition of the factor structure. The same participants completed the Short Form Health Survey for construct validation. MuRQoL responses from 68 CI users were compared with those of a matched group of normal-hearing (NH) adults. **Results:** Eighteen items measuring music perception & engagement and 18 items measuring their importance were selected; they grouped together into two domains. The final questionnaire has high internal consistency and repeatability. Significant differences between CI users and NH adults and a correlation between music engagement and quality of life (QoL) support construct validity. Scores of music perception & engagement and importance for the 18 items can be combined to assess the impact of music on the QoL. **Conclusion:** The MuRQoL questionnaire is a reliable and

valid measure of self-reported music perception, engagement and their importance for adult CI users with potential to guide music aural rehabilitation.

Keywords: cochlear implant, quality of life, impact, questionnaire, validity, reliability

Abbreviations: CI: cochlear implant, ICC: intra-class correlation coefficient, MuRQoL: music-related quality of life, NH: normal-hearing, QoL: quality of life, SF12v2: 12-item short form health survey version 2

Introduction

Cochlear implant (CI) users do not receive much of the temporal fine structure necessary for accurate perception of pitch and timbre; as a result, many CI users are dissatisfied by the music they hear (Looi et al., 2012). However, new music-focused CI technologies and auditory music training programmes may have the potential to benefit certain aspects of music perception and enjoyment with potential subsequent benefits for quality of life[1] (QoL) (van Besouw et al., 2015; Limb & Roy, 2013). Demand for music training is high; in a study by Looi & She (2010) 46 out of 84 adult CI users were willing to participate in a music training programme. Evidence for the effectiveness or not of these interventions is needed to enable patients, clinicians and other stakeholders to make informed decisions regarding how much time and money, if any, they should invest in music rehabilitation. Although music outcomes have been traditionally measured using music perception tests, their scores do not necessarily predict music appraisal, enjoyment and participation in musical activities (Drennan et al., 2014; Wright & Uchanski, 2012). Wright & Uchanski (2012) found weak, if

[1] "An individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns" (WHOQOL, 1993)

any, correlations between music appraisal ratings and music perception scores, which suggests that music enjoyment should not be assumed, but explicitly studied in conjunction with music perceptual accuracy.

Need of a new self-report measure for music

Self-report measures can better capture the wider impact of interventions for CI users than perception tests that assess specific auditory abilities. However, existing music questionnaires designed for adult CI users, i.e. the Iowa Musical Background Questionnaire (IMBQ), the Munich Music (MUMU) questionnaire and the University of Canterbury Music Listening Questionnaire (UCMLQ) were not designed for measuring rehabilitation outcomes or intended as standardised tools for assessing CI users' music experiences (Gfeller et al., 2000; Looi & She, 2010; Brockmeier et al., 2002). The IMBQ was designed to describe changes in music experiences (formal music training, time spent listening to music) pre- to post-implantation, the UCMLQ to inform the development of a music training programme and the MUMU to compare the effects of various processing strategies on music listening habits (e.g. musical instrument playing and singing). Aspects of the relationship of CI users with music that are relevant to the QoL, such as feelings about music or music-related social interaction, are also poorly covered in previous questionnaires, which limits the range of music-related benefits of music rehabilitation they can detect. None of these questionnaires uses a uniform response scale, which makes them difficult to score; the response options of the MUMU and the IMBQ have had to be adapted (to a 5-point Likert scale for the MUMU and a 3-point scale, positive, neutral and negative for the IMBQ) to facilitate statistical analysis of responses (Veekmans et al., 2009; Drennan et al., 2014). Finally, although the content and face validity of the IMBQ, the MUMU and the UCMLQ have been assessed, there is no

published evidence for their psychometric properties, reliability or construct validity, which are required for a questionnaire to be a strong measure.

In addition to patient-specific factors (e.g. residual hearing) and device-specific factors (e.g. implant type), CI users' music appraisal (i.e. liking), enjoyment and activity might also depend on expectations, musical background and preferences. Some CI users listen to music less after implantation because they are disappointed by the music sound quality in contrast to how they expect it to sound, either based on what they think normal-hearing (NH) people would hear or, in the case of postlingually deafened adults, based on what they used to hear (Gfeller et al., 2008; Bartel et al., 2011; Migirov et al., 2009). Expectations are likely to have an impact on the role of music in life and its effect on the QoL. Zhao et al. (2008) and Lassaletta et al. (2008) showed a positive correlation of music sound quality and enjoyment with QoL in postlingually deafened adult CI users (Zhao et al., 2008; Lassaletta et al., 2008). Also Calvino et al. (2015) found a positive correlation between the music perception scores of the Hearing Implant Sound Quality Index and the overall score of the Glasgow Benefit Inventory (Calvino et al., 2015). However, Fuller et al. (2013) did not find a similar association for a group of prelingually deaf and late implanted adult CI users, possibly due to the different expectations of this group as a result of limited NH memory of music (Fuller et al., 2013). It is crucial to assess what is important for the individual and the relative impact of music on their QoL, not only when assessing benefits from music rehabilitation, but also for the identification of individual rehabilitation needs and monitoring individual patients in clinic. A valid and reliable psychometric instrument measuring music experiences and their impact on the QoL is needed to guide and evaluate music rehabilitation tools for adult CI users.

The prototype Music-Related Quality of Life (MuRQoL) questionnaire

The MuRQoL concept was initially defined as the impact of CI users' music listening abilities, attitudes and activities on their QoL, which is a function of their perceived importance. On the basis of this concept, focus groups with adult CI users were used for the generation of items for the questionnaire, to ensure its content validity. This approach differs from previous studies, where CI users were asked to comment on questions but were not involved in the initial process of item generation (Gfeller et al., 2000; Brockmeier et al., 2002). Fifty-three items resulted from six focus groups with 30 adult CI users of a wide age range and musical background, both prelingually deaf and postlingually deafened (Dritsakis et al., in press). Items were identified using 'template analysis' of the focus group data. The prototype MuRQoL questionnaire addressed novel aspects of music experience, such as the ability to hear whether you are singing in tune and used a 5-point frequency Likert scale ranging from 'Never' to 'Always'. The present study is concerned with the refinement of the prototype MuRQoL questionnaire, psychometric item selection and validation.

Psychometric properties

Reliability refers to the homogeneity of a questionnaire (internal consistency) and to its ability to produce the same results in subsequent administrations when no clinical change happens in the meantime (test-retest reliability). These ensure that all items in a questionnaire tap the same concept and that a detected change can be attributed to the given condition, respectively. Reliability criteria that are required for the assessment of individual patients are higher than those needed for group measurements (McHorney & Tarlov, 1995). Validity refers to the extent that a questionnaire assesses what it is intended to assess. Content validity refers to the relevance of the items to the users, whereas face validity refers to their clarity

and appropriateness; they are both assessed qualitatively with the intended users or with experts. The face validity of existing music questionnaires for adult CI users has been commonly assessed with expert input (Gfeller et al., 2000). Construct validity assesses whether a questionnaire measures the theoretical construct of interest. It has been tested through the 'known-group' or extreme group method, whereby a measure is valid if it gives significantly different scores for two groups that are known to be different on a specific concept (Yang et al., 2013). Another method for construct validation is to check if a questionnaire correlates with another that measures the same concept (convergent validity). Additional measurement properties required for individual items are discriminability, minimum floor/ceiling effects and minimum inter-item correlations. Use of psychometric techniques for item selection ensures high measurement ability for both the items and questionnaire. Psychometric techniques have not been used for item selection or validation in existing music questionnaires with adult CI users, but they are common in the development of other CI-specific questionnaires (Amann & Anderson, 2014; Hinderink et al., 2000).

Aims of the present work

This paper describes the refinement and validation of the MuRQoL questionnaire across four studies. Study 1 aimed at improving the wording of the 53 prototype items to ensure face validity, removing redundant items and improving the domain structure. The aims of Study 2 were to select the items with the best psychometric properties, to provide evidence for the questionnaire's reliability and to establish its factor structure. The aims of studies 3 and 4 were to assess the construct validity of the questionnaire. Study 3 used the 'known group' method by comparing CI users' scores with the scores of NH adults and Study 4 assessed convergent validity through examining whether the scores of the MuRQoL questionnaire

correlate with the scores of a generic QoL instrument. All studies were approved by the University of Southampton Research Governance Office and Faculty of Engineering and the Environment Ethics Committee (8264) and the UK National Research Ethics Service (14/EM/0140).

Study 1: Expert review

Methods

Twenty-four adult professionals (9 male, 15 female, mean age: 39.5 years old) with a wide range of expertise relevant to CIs or music (10 researchers, 7 audiologists, 2 speech and language therapists, 3 ENT doctors, 1 music therapist and 1 composer) participated in an online survey. They were recruited through personal invitations, the AUDITORY mailing list, UK CI centres, CI manufacturers and the British Association of Music Therapy over a period of two months. Participants were given the 53 prototype questionnaire items in a random order and were asked to comment on the clarity and appropriateness of each item. Repeated comments were used for the improvement of the wording of the items. A refined MuRQoL questionnaire version 2 (MuRQoLv2) was developed.

Results

Items that were not considered clear or appropriate were modified; for example, in the question ‘Can you hear differences in musical pitch?’, ‘pitch’ was replaced by the less technial description ‘musical tone, i.e. how high or low music is’. The ambiguity of items was minimised; for example, in the question ‘Can you hear the words in music’, ‘hear’ was replaced by ‘recognise’ to avoid confusion with sound detection. Items that were overlapping were merged with others, reducing the total number of items in the MuRQoL questionnaire

from 53 to 46. These changes led to the MuRQoLv2. For each question, the same 5-point Likert scale was used: Never, Rarely, Occasionally, Frequently, Always. This scale has been previously used in the Musical Stages Profile, a music questionnaire for paediatric CI users, and in the Cochlear Implant Function Index, a measure of auditory abilities of adult CI users (Coelho et al., 2009; Edwards, 2014).

For each of the 46 items assessing music listening abilities, attitudes and activities on the 5-point frequency Likert scale (hereafter, 'frequency items' or 'frequency scale[2]'), a corresponding item was developed to gauge its importance to the respondent, using the 5-point Likert scale: Not important at all, Not very important, Somewhat important, Very important, Extremely important. For example, for the MuRQoLv2 frequency item 'Can you hear the beat in music?' the corresponding importance item was 'How important is it for you to be able to hear the beat in music?' These items will be referred to from now on as 'importance items' or 'importance scale'. The assessment of the importance of music was in line with the definition of the MuRQoL concept, whereby the impact of music on the QoL is a function of musical ability, attitude or activity and their importance. The need to assess importance was supported by strong differences between CI users with regards to the role of music in their life in the focus groups (Dritsakis et al., in press). Responses to the importance questions could help weight the responses to the frequency questions to measure the impact of music on the QoL. This would mean that the impact of, for example, the ability to recognise familiar music, on the QoL would be different depending on how important this task would be for an individual (see 'Discussion'). Use of importance ratings to produce QoL scores has been previously used in the Quality of Life Profile (Renwick et al., 2003). In the

[2] The term 'scale' here refers to a set of items and is different than the term 'Likert scale' which refers to the response options.

present work, the importance scores were used to remove relatively unimportant items (Study 2) and for construct validation (Study 4).

Study 2: Psychometric item selection

Methods

One hundred forty-seven adult CI users (58 male, 89 female, mean age: 56.69 years old, mean duration of CI use: 6 years and 9 months) were recruited through the University of Southampton Auditory Implant Service, the St. Thomas Hearing Implant Centre (London), the Ear Foundation and Action on Hearing Loss over four months. Thirty-six out of 147 CI users were prelingually deaf and ten had received some formal music training (Table 1). They completed the 46-item MuRQoLv2 that was developed in Study 1 (46 frequency and 46 importance questions) online or by post. A not applicable (N/A) option was included with the response options to determine which questions were applicable to the majority of participants and therefore most useful to retain in the final version of the questionnaire. Both frequency and importance questions were presented to participants under the three domain headings: *ability, attitude* and *activity*. One hundred thirty-three participants repeated the questionnaire two weeks later. The 2-week gap was considered long enough for respondents not to remember their answers and short enough for there to be no changes to participants' music experiences. Participants were asked to report on any clinical change (which could have affected their music experiences) prior to completing the questionnaire the second time[3]. Psychometric techniques of the classical test theory were used: analysis of items (distribution of responses per item, test-retest reliability of each item, inter-item correlations), assessment

[3] Eight participants reported clinical changes (e.g. processor upgrade) between the two administrations but test-retest analysis with and without these participants suggested that even if there was a substantial test-retest change in their music experience, it did not affect the repeatability of the items. Another participant informally reported an important change in their attitude towards music (meeting with musical family, having ukulele lessons) and was excluded from the test-retest analysis.

of internal consistency and factor analysis (Streiner et al., 2015: chapter 5). Response options corresponded to 1-5 numerical values; for the eight negatively phrased frequency items, scores were reversed. The Likert-scale data were treated as interval, which is a common convention in questionnaire development (Hinderink et al. 2000; Amann & Anderson 2014; Streiner et al. 2015: 52-53). The percentage of N/A responses was calculated for each frequency and importance item, and the range and distribution of scores were calculated for each frequency item. Correlations between frequency items were examined and items that correlated very highly or very poorly with other items were considered for removal (Field 2014: 685-686). Test-retest reliability was measured using the 'intra-class correlation coefficient' (ICC), which measures score changes over time by comparing differences within subjects with the total variance (Deyo et al., 1991). If individual subjects give similar judgements across time, the total variance is dominated by inter-subject variability and the ICC is high. Values that have been recommended for the interpretation of kappa or weighted kappa coefficients were adapted here for ICCs (Landis & Koch, 1977). Exploratory factor analysis was performed for the frequency items to identify underlying factors within the questionnaire and items that group together in each of these factors (Field 2014: chapter 17). Cronbach's α correlation coefficient was calculated to assess the internal consistency of the questionnaire and its domains and to detect items that reduce homogeneity.

[Table 1 around here]

Results

Twenty-five frequency items and their corresponding importance items were removed based on descriptive statistics, test-retest reliability and inter-item correlations. Three items were

removed because they were rated ‘Not important at all’ or ‘Not very important’ by > 60% of the participants and another three items due to > 10% ‘N/A’ responses for frequency or importance. Three other items were removed because they were scored ‘Frequently’ or ‘Always’ by > 60% of the participants. Elimination of items with low overall importance and items with floor or ceiling effects ensured that only the items that matter for CI users and those that can discriminate well between individuals were retained for the final version. The question ‘Can you recognise music that you have not heard before?’ was removed because it was considered confusing by participants. The question ‘Do you make the effort to listen if the music is hard to recognise or follow?’ was removed because of potential overlap with the item ‘Can you hear music without effort or having to concentrate?’, according to the subjective judgement of the first author. ICCs were calculated for each frequency and importance item using a 2-way mixed effects model for absolute agreement (single measures). Analysis indicated seven frequency items with repeatability < 0.2, which were removed. Ten items were removed due to very high (> 0.7) or weak (< 0.3) Spearman correlations with other items. All the remaining 18 frequency and 18 importance items had a standard deviation (SD) of >1 for frequency or importance, while 15 had a SD >1 for both and all their response options were used.

Exploratory factor analysis was performed on the remaining 18 frequency items using Principal Component Analysis (PCA) for factor extraction with oblique rotation using SPSS version 22 (IBM Corp., 2013). PCA identifies factors that account for the variance in the data; each factor has an eigenvalue, which represents the power of the factor to account for the variance. Eigenvalues > 1 are used by convention to determine factors. Rotation facilitates interpretation of the factors (oblique rotation was chosen because the factors were unlikely to be completely independent). Each item on a factor has a loading, which shows the correlation between the factor and the items. Factor loadings are interpreted relatively to each

other, i.e. an item with a higher loading than another on a factor is interpreted as fitting better to that factor. A value of 0.945 for the Kaiser-Meyer-Olkin test of sampling adequacy was within the range of 'marvellous' and suggested that the sample was adequate for the analysis to give reliable results (Field 2014: 17.7.1). Analysis showed two factors: factor 1 (eigenvalue = 8.9, explained 49% of the variance) comprised 13 out of 18 items and was interpreted as a 'music perception' domain. Nine of the 13 items were initially under the *ability* domain (Study 2). The remaining five items (originally in the *activity* domain) clustered together in factor 2 (eigenvalue = 5.6, explained 31% of the variance), which was interpreted as a 'music engagement' domain. The two enjoyment items ('Do you enjoy music in noisy environments when no visual cues are available (e.g. at a party, at a restaurant or in the car over the engine/road noise)?' and 'Do you enjoy music on TV, DVD or on the computer when visual cues are available?' had similar loadings on both factors and were grouped under factor 2. The inclusion of music enjoyment and activity within the engagement domain is in agreement with previous studies, whereby attitude towards music and participation have been grouped together under the concept of music engagement (Gfeller et al., 2012). The two domains are referred to from here on as 'perception' and 'engagement' subscales.

Calculation of Cronbach's α indicated excellent (> 0.80) internal consistency for the frequency and importance scales and for each of the perception and engagement subscales (Table 3). The overall frequency scale and perception subscale had strongest Cronbach's α and narrowest confidence intervals. No items were found to significantly increase internal consistency if deleted so no more items were removed.

The final version of the MuRQoL questionnaire after item selection (MuRQoLv3, Appendix 1) therefore consisted of 18 frequency items (frequency scale, Never – Always) grouping into the two meaningful subscales: 'music perception' and 'music engagement' and 18

corresponding importance items (importance scale, Not important at all – Extremely important).

The ICCs for the overall frequency and importance scales and subscales showed excellent test-retest reliability (> 0.80) and were higher for perception than engagement (Table 2). Not only were the point estimates for the ICC values > 0.80 but the lower bounds of their confidence intervals were also above this value (except for frequency engagement), giving a high degree of certainty that the ICC is sufficiently high to justify the use of these scores at the group level, and at the individual level for the overall frequency scale (and perception subscale), where the lower bound for the ICC was > 0.9. The same table summarises the mean score change (M_{change}) at retest and the associated SD for the frequency and importance scores and subscale scores. The minimum change that was assumed to represent a 'true' change or the 'smallest detectable change' (SDC) for each of the overall scales and subscales was calculated based on the following formula: $1.96*SD_{change.}$ Where the M_{change} was significantly different from zero, it was added to the formula: $M_{change} + 1.96*SD_{change}$ (Guyatt et al., 1987). SDCs for all scores are reported in Table 2.

[Table 2 around here]

Study 3: Comparison with NH adults

Methods

Sixty-eight adults without known hearing difficulties[4] (NH group) (48 male, 94 female, mean age: 37 years old, age range: 18-87) completed the 46-item MuRQoLv2 in the form of an

[4] It is acknowledged that some participants may have some hearing loss of which they are unaware.

online survey. Participants were recruited through the University of Southampton over three months and were students, technicians, administrators, academics and other university staff. As with the CI sample, it is likely that people with an interest in music were keener to participate.

Sixty-eight of the 147 adult CI users from study 2 were selected, who matched the 68 NH adults for age, gender and professional music training (Table 3). Age and prior music experience were previously used as matching criteria by Veekmans et al. (2009). The 68 selected CI users had a mean duration of CI use of 6 years and 6 months; 27 of them were prelingually deaf (i.e. congenitally deaf at least in one ear or went deaf < 3 years old), 39 were unilateral, 24 were bimodal and five were bilateral CI users. It is unclear to what extent these samples are representative of the general CI and NH populations, given the potential confounding effects of other variables such as socioeconomic status or education level.

Only responses to the selected final 18 frequency and 18 importance items were used for both CI users and NH adults. Average scores for the frequency and importance scales and perception and engagement subscales were calculated for each participant of the CI and NH group. It was hypothesised that the MuRQoL frequency scores of NH adults would be significantly higher than those of CI users for the overall scale and subscales. This hypothesis was based on the evidence in the literature for significantly worse performance of adult CI users compared to NH adults in the perception of pitch, timbre and in the recognition of familiar melodies, as measured with music perception tests (Kang et al., 2009; Looi et al., 2008; Brockmeier et al., 2011). Significantly poorer scores for self-reported music enjoyment and music listening habits have also been reported for CI users (Veekmans et al., 2009). No hypothesis was made regarding the size of the difference because the MuRQoL questionnaire is a new measure and the study was a first step towards exploring its measurement properties.

No hypothesis was made for the importance scores due to the lack of evidence regarding differences in the importance of music between CI users and NH adults.

Results

For all frequency scores (overall scale and perception and engagement subscales) the NH group had statistically significantly higher scores than the CI group, as indicated by the results of the Mann-Whitney test (Table 4). This effect was stronger for perception and for the overall frequency score than for engagement. The largest median difference between the two groups was in the range of one Likert-scale point and specifically 1.22 for the engagement subscale of the frequency scale (frequency-engagement). No statistically significant difference was found for importance; the small differences between the median scores of the two groups for the perception subscale and the overall frequency scale were in favour of the CI group. Figure 1 shows the overall frequency and importance score distribution across groups. There was large variability in the CI group with average frequency and importance scores covering the whole Likert scale and overlapping with the NH responses (see Figure 1 and the range of scores in Table 4). Multiple regression showed that all the background variables together (age, gender, music training, type of deafness, CI configuration, duration of CI users) accounted for only about 11% of the total response variance for frequency (R^2:0.119) and importance (R^2:0.110) scores. There was a significant effect of age on frequency (standardised beta: -0.310, p=0.002) and music training on importance (standardised beta: 0.193, p=0.020), but both effects were weak.

[Table 4 around here]

[Figure 1 around here]

Study 4: Correlation with SF12v2

Methods

A QoL measure was employed for the assessment of convergent validity, due to the lack of validity of previous music questionnaires designed for adult CI users. For individuals who rated music as important, frequency scores for music perception and engagement were anticipated to positively correlate with the QoL scores. This would be evidence for the validity of the MuRQoL construct because it would suggest that the combination of frequency and the importance scores can show the impact of music on CI users' QoL. The 147 CI users who completed the 46-item MuRQoLv2 in Study 2 also completed the 12-item Short-form Health Survey version 2 (SF12v2) on paper or online. The SF12v2 is a QoL questionnaire that can be used to produce utility scores for economic evaluations (Ware et al., 1996). It gives a 0-100 score for each of eight domains (physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional and mental health), two average scores, i.e. the Physical Component Summary (PCS) and the Mental Component Summary (MCS), as well as a single utility score (SF6D). It is a shorter version of the SF36, it covers the same eight domains and closely replicates the SF36 summary scores (Ware et al., 1996). SF36 has been shown to be sensitive to the benefits of CIs in the domains 'general health', 'mental health', 'physical role functioning', 'emotional role functioning', 'social functioning' and the 'MCS' (Loeffler et al., 2010). The SF12v2 has improved wording and scoring compared to version 1. The summary scores of version 1 have been shown test-retest reliable (ICCs= 0.77 and 0.86) in the general UK population (Ware et al., 1996).

CI users' responses to the 18 final MuRQoL items from Study 2 were used. The MuRQoLv2 overall frequency score and scores for the perception and engagement subscales were calculated. Correlations between MuRQoLv2 frequency scores and SF12v2 scores were

expected only for participants who rated music as at least 'Somewhat important' on average, based on the assumption that music would have an impact on the QoL only when it is considered important. Correlations were expected between the MuRQoL frequency scores for the perception & engagement subscales and the SF12v2 PCS & MCS scores, as well as between the MuRQoL overall frequency score and the SF6D utility score. This was based on the evidence in the literature for an association between self-reported music perception, enjoyment and sound quality and the QoL of CI users, as well as the evidence for an impact of music on the physical, psychological and social aspects of the QoL of adults without hearing problems (Schäfer et al., 2013; Chin & Rickard, 2012).

Results

A positive correlation was found between the MuRQoL frequency-engagement score and the SF12v2 PCS for an importance score of ≥ 3 for engagement [Spearman r=.229, p<0.5, n=91] (Figure 2). No other correlation was significant. Among the four domain scores contributing more to the PCS there was a significant correlation between the MuRQoL frequency-engagement and the SF12v2 RP score for an importance score ≥3 (Spearman r: .260 (p<.05) [N=91]), addressing activity limitations (at work or any other daily activity) due to physical health (Spearman r=.260, p<.05). This correlation shows that to some extent CI users who rated their enjoyment of music or their participation in musical activities highly, reported fewer activity limitations. When importance scores for the engagement subscale were < 3, the correlation between frequency-engagement score and the SF12v2 RP (r=.083, p=.548, N=55) was not significant, which supports the prediction that music has an impact on the QoL only when it is considered important by an individual.

[Figure 2 around here]

Discussion

This research is novel in using psychometric methods for the item selection and validation of a music questionnaire for adult CI users. In previous questionnaires item selection was only based on expertise and content or face validation (Gfeller et al., 2000; Looi & She, 2010; Brockmeier et al., 2002). Psychometric item selection with the use of CI data ensured face validity, discriminability, test-retest reliability, minimum overlap between items and floor and ceiling effects of the final 18 frequency and importance items. It also ensured the test-retest reliability, internal consistency and factor structure of the overall frequency scale and its subscales. The face validity of the final 18 frequency items was also supported by the expert review. The high SD of the items and the wide range of scores suggest that items can discriminate among CI users with varying levels of self-perceived music perception, music engagement and their importance. The elimination of items correlating highly with others ensured the minimum number of questions with differentiating content to be able to capture a breadth of music experiences (McHorney & Tarlov, 1995). The use of one uniform Likert scale for 'frequency' and one for 'importance' facilitated calculation of average scores for overall frequency and importance scales and for the perception and engagement subscales. The findings of the present studies suggest the high measurement properties of the MuRQoLv3 questionnaire, in addition to the strong content validity supported by the involvement of CI users in the item generation phase (Dritsakis et al., in press). Five of the 18 final frequency items are also novel among music questionnaires designed for adult CI users, namely items 4, 6, 9, 10 and 15 (see Appendix 1); this is important for capturing music experiences that have not been measured with previous questionnaires. A number of items are similar to items from previous music questionnaires designed for adult CI users, namely the

IMBQ (e.g. item 12), the MUMU (e.g. item 18) and the UCMLQ (item 5), which shows that the MuRQoL builds on previous instruments.

From a theoretical point of view, this was the first study to use psychometric techniques to group the music experiences of CI users into domains. The final 2-dimensional distinction is generally in line with the traditional distinction between 'music perception' and all aspects not related to perception, commonly referred to as 'music appreciation' (Looi et al., 2012). However, the music sound quality question 'Does music sound in tune?' clustered together with music perception items, in contrast with the traditional approach in the CI literature whereby music sound quality attributes have been assessed together with enjoyment and music listening habits. The question was likely to have been perceived by the respondents as a perception item, i.e. 'Can you hear whether music is in tune?' or 'Can you hear music in tune?'. This has implications for the assessment of music sound quality and the phrasing of music sound quality questions.

Reliability

The test-retest reliability (indicated by ICC) for all the frequency measures and the Cronbach's α for the overall frequency and importance scales and perception domain subscales exceed the 0.90 criterion that has been recommended for individual level measurements (McHorney & Tarlov, 1995). This, in addition to the evidence for validity and the measurement properties of the individual items, suggests that the MuRQoLv3 could be a useful measure not only for the assessment of music experiences at the group level but also for diagnosing and monitoring the difficulties of individual patients with music in clinic. However, which changes in music experiences are clinically important has to be assessed with the use of clinical anchors at the individual level (Revicki et al., 2008). Given the lack of

a validated clinical music perception test, clinician ratings post-intervention could be used to identify individuals who have changed; MuRQoL change scores in these individuals could be interpreted as clinically important change.

The < 0.90 Cronbach's α and wider confidence intervals for the frequency and importance engagement subscales may be related to the fewer number of items (seven) in this subscale and could increase by increasing the number of items. However, item selection was based on other criteria that ensured the measurement properties of the individual items. Yet, the internal consistency was very high and appropriate for group measurements. For individual measurements in clinic, the overall scores can be used to maximise internal consistency. The other psychometric properties of the engagement subscale (repeatability, statistically significant difference between the CI and NH group, correlation with the SF12v2) and its importance for content validity support its usefulness. The < 0.90 test-retest reliability for all importance scores suggests that the importance scores may not be appropriate to reliably assess individual patients in clinic. The lower repeatability for the importance scale and subscales suggests that participants found these items difficult to answer; this is further supported by participants' anecdotal comments (e.g. 'How important is it for me to hear the beat, rhythm or instruments, voices etc.? My answer is 'Extremely important'. But as I can't hear any of these things it also has to be N/A'). Despite these limitations, the repeatability of the importance measures (> .80) is suitable for group assessments (Nunnally & Bernstein, 1994: 264-265).

To the knowledge of the authors, the sample of 147 adult CI users is the largest sample that has been reported in the literature for the pilot-testing of questionnaires designed for CI users (Brockmeier et al., 2002; Hinderink et al., 2000; Amann & Anderson, 2014). The MuRQoL questionnaire was developed and validated with and was applicable to a wide range of adult CI users (prelingually deaf and postlingually deafened, musically trained and non-musical)

which suggests that it has potential to be used by the majority of the adult CI user population. However, anecdotal comments and the percentage of N/A responses indicate that the questionnaire may not be equally informative for everyone. CI users with limited exposure to music post-implant, such as the prelingually deaf, considered many of the questions (especially the importance questions) irrelevant due to their difficulty in understanding or appreciating music in any way. For example, the question 'Does music sound in tune?' had more N/A responses than any other item (five) and only by prelingually deaf participants, probably due to never having had a 'normal' perception of music.

Construct validity

A combination of techniques was used for the construct validation of the MuRQoL questionnaire. The statistically significant difference in the frequency scores for music perception and engagement between CI users and their NH peers shows that the MuRQoL questionnaire can detect differences between two groups which are known to differ in their music experiences. The result is in line with the previous findings of Veekmans et al. (2009), who reported significantly better performance in some items of the MUMU for NH adults than for unilateral CI users and generally better scores than bilateral CI users (Veekmans et al., 2009). The NH range and average scores can be used as normative data to interpret CI scores in future studies. However, it is important that longitudinal assessment establishes a personal norm for each patient based on which they should be compared.

The one Likert-scale point average difference between the CI and NH group was above the smallest detectable change (SDC) found for the overall frequency score (4% or 0.2 Likert scale point). The very small and non-significant differences for importance indicate that CI users consider music equally important as their NH peers, despite poorer self-reported

perception and engagement. This finding supports the need to improve the music perception and enjoyment of CI users, with potential subsequent benefits for the QoL. It is also in agreement with previous findings. Frederigue-Lopes et al. (2014) adapted the MUMU into Brazilian Portuguese and used it with 19 postlingually deafened CI users. They found a decrease in listening habits post-implantation, but music was still rated as important (Frederigue-Lopes et al., 2015). The large variability in the frequency scores for the CI users, with many CI users scoring similarly to NH adults, is in agreement with previous studies, which show high variability in CI outcomes for music perception and self-perceived music experiences (Maarefvand et al., 2013; Gfeller et al., 2008). The high variability in average CI scores suggests that CI users may or may not require support with music; this way, the variability stresses the importance of assessing individual needs.

The moderate positive correlation between the MuRQoL frequency-engagement and the SF12v2 RP domain (which covers activity limitations as a result of physical health) for participants who find music important, shows that the MuRQoL questionnaire can predict some aspects of QoL. An interpretation of this correlation is that music enjoyment and participation in musical activities have an impact on general social activity. For example, poor music enjoyment and limited musical activity may prevent CI users from accomplishing as much as they would wish to in their everyday life. For those with a music-related occupation, this may reflect limitations in fulfilling their professional duties. However, it is not clear if this association suggests an impact of music enjoyment or activities on general social functioning or vice versa. Yet, the findings are in agreement with previously reported correlations between music enjoyment or sound quality and QoL in CI users (Calvino et al., 2015). The lack of any other association between the MuRQoL subscale scores and SF12v2 domain scores may mean either that the SF12v2 covers aspects of CI users' health that are

not related to music or that, although music has an impact on the QoL, still other components of QoL are more important than music and affect QoL more.

Combination of frequency and importance scores

The assessment of the importance of music experiences and the potential of the MuRQoL questionnaire to measure the impact of music on the QoL overcome limitations of previous approaches that either did not address the relationship between music and the QoL or used correlations between music enjoyment or appraisal and QoL that were measured separately (Calvino et al., 2015). The frequency and corresponding importance scores for each item can be plotted across each other to produce a matrix showing the impact of each of the 18 music perception/engagement tasks on the QoL of individual CI users (Table 5). When music is rated as important, but the frequency score for the perception or engagement subscale (or the overall frequency scale score) is poor, music is expected to have a strong and negative impact on the QoL (shaded area in the matrix). A patient scoring in this 'critical block' for a music perception/engagement item indicates a need for music rehabilitation and the specific aspects of music to be targeted (e.g. pitch perception or singing). In this way the MuRQoL measure could be used as a diagnostic tool to identify individuals who need support with music and to guide music rehabilitation.

[Table 5 around here]

Recommendations for use

To identify rehabilitation needs in clinic, the scores for the 18 final frequency and importance items can be plotted on a matrix, as shown in Table 5. The number of items falling in the

critical block of the matrix for an individual patient can highlight areas of concern and can help forming a profile for that individual. The MuRQoL questionnaire can also be used to measure changes in music experiences post-intervention. For group comparisons it is recommended that the overall frequency and importance scores and/or the perception and engagement subscale scores are used depending on the aims of the study. For example, if one is assessing the effects of musical instrument recognition training on music engagement in particular, the average frequency scores of the engagement subscale could be used. If the interest is in the effects of a CI on the importance of music, the average overall frequency score could be used. For individual patients it is recommended that only the frequency scores are used because only these exceeded the 0.90 test-retest reliability criterion for individual measurements. The NH range and average scores can be used as normative data to interpret CI scores in future studies.

For future use, changes are recommended in the phrasing of some frequency items and their corresponding importance items due to the elimination of similarly phrased questions in Study 2 (see Appendix 2 for the recommended changes). These changes are included in the MuRQoLv4, which is available for use; see 'Data Access Statement'.

Future work

The responsiveness of the items of the MuRQoL questionnaire to changes after music interventions should be assessed in a future longitudinal study. The minimum significant change scores calculated here can be used for the interpretation of future results, but it is important that clinically meaningful changes are determined with the use of clinical anchors or individual interviews with the patients to confirm that detected changes are associated with a true change in music experience. It is recommended that clinical changes on the individual

level are measured only using the overall frequency score and frequency subscale scores, which fulfilled the recommended 0.90 reliability criterion for individual measurements. A follow-up study could also examine correlations between CI users' MuRQoL frequency-perception scores and their scores in music perception tests. This would help professionals gain a better understanding of the relationship between the patient's self-perception and perceptual accuracy.

Conclusions

The MuRQoL questionnaire consists of two sets of 18 items each, one assessing music experiences and another assessing their importance. The items have content and face validity, and each set of items is measured on a uniform 5-point Likert scale; five of the items are novel among music questionnaire designed for adult CI users. The items grouped together into two meaningful domains with high internal consistency and test-retest reliability. Significant differences between adult CI users and NH adults, as well as a correlation between the MuRQoL questionnaire and QoL scores were evidence for construct validity. In this way, the MuRQoL questionnaire overcomes weaknesses of previous music questionnaires designed for adult CI users and has the potential to fill the gap for a reliable and valid outcome measure for the evaluation of music-focused interventions. However, a future longitudinal study should assess the ability of the questionnaire to detect clinical changes. The MuRQoL questionnaire also has the potential to be used as a screening tool to identify individual rehabilitation needs in clinic through the measurement of the impact of music on the QoL. The clinical utility of the MuRQoL questionnaire is supported by the strong reliability, which exceeds the recommended criteria for individual measurements. However, the clinical usefulness of the questionnaire should be demonstrated with clinical use and experience.

Tables

Table 1 Characteristics of the CI users (n=147) who participated in studies 2 and 4.

Age (years)	18-84, mean age = 56.69, SD = 16.02
Duration of CI use	4 months - 26 years, Mean = 6 years and 9 months, SD = 6 years, 8 months
Gender	58 male, 89 female
Type of deafness	109 postlingually deafened, 36 prelingually deaf
Music training	10 had received formal music training, 137 had not
CI configuration	87 unilateral, 9 bilateral, 47 bimodal, 3 users of electro-acoustic simulation (EAS)
Type of administration	135 completed the questionnaire online, 17 by post
CI manufacturer	75 Cochlear Ltd., 36 Advanced Bionics, 30 MED-EL, 2 Neurelec and 3 did not define

Table 2. MuRQoL questionnaire internal consistency, test-retest reliability, mean score change at retest (M_{change}) and smallest detectable change (SDC). SDC was calculated as $1.96*SD_{change}$ for the frequency change scores and as $M_{change} + 1.96*SD_{change}$ for the importance change scores, which were significantly different from 0 (Guyatt et al., 1987).

Measure	**Cronbach's α (n=147)**	**ICC (n=133)**	**Mean score change (SD)**	**Smallest detectable change****
Frequency OVERALL	.947 (.932 - .959*)	.964 (.950 - .974*)	-0.2% (4.18%)	8%
Frequency *perception*	.940 (.923 - .954)	.941 (.918 - .958)	-0.32% (5.82%)	11%
Frequency *engagement*	.840 (.796 - .878)	.951 (.931 - .965)	-0.08% (5.26%)	10%
Importance OVERALL	.920 (.898 - .939)	.850 (.795 - .891)	0.9% (7.68%)	15%
Importance *perception*	.904 (.878 - .926)	.802 (.732 - .855)	0.84% (9.12%)	18%
Importance *engagement*	.858 (.818 - .893)	.842 (.785- .885)	1.08% (9.28%)	19%

*95% confidence intervals; **rounded to nearest integer

Table 3. Characteristics of the matched groups of CI users and NH adults employed in Study 3.

	CI users (n=68)	**NH adults (n=68)**
Age range	18-80 years	
Mean age	45.9 years	45.8 years
SD age	15 years	
Gender	51 females and 17 males	49 females and 19 males
Music training	3 out of 68 had received formal music training	

Table 4. Results of the comparison between the MuRQoL scores of the CI and NH group. The threshold for a large effect is set at: 0.5.

Measure	Median (range)	Mann-Whitney
Frequency TOTAL	CI: 3.17 (1 – 4.67) NH: 4.06 (2.89 - 4.67)	U=586.5 (p<0.001) R= -0.65
Frequency *perception*	CI: 3.14 (1 – 4.73) NH: 4.36 (3.45 - 5)	U=421.5 (p<0.001) R= -0.71
Frequency *engagement*	CI: 3.14 (1 – 4.57) NH: 3.60 (1.86 – 4.71)	U=1235.5 (p<0.001) R= -0.40
Importance TOTAL	CI: 3.61 (1.22 – 4.89) NH: 3.50 (2 – 4.56)	U=2474.5 (p=0.478) R= 0.06
Importance *perception*	CI: 3.73 (1.27 – 4.82) NH: 3.64 (1.91 – 4.82)	U=2491 (p=0.435) R= 0.07
Importance *engagement*	CI: 3.43 (1.14 - 5) NH: 3.43 (1.29 – 4.29)	U=2414.5 (p=0.655) R= 0.04

Table 5. Matrix of frequency and importance score combinations. The shaded area indicates scores that could be interpreted as 'critical' for clinical rehabilitative purposes.

<table>
<tr><th></th><th colspan="3">1 (NEVER) 2 (RARELY) 3 (OCCASIONALLY)</th><th colspan="2">4 (FREQUENTLY) 5 (ALWAYS)</th></tr>
<tr><td>1 (NOT IMPORTANT AT ALL)</td><td colspan="3" rowspan="2">POOR IMPORTANCE & POOR FREQUENCY
WEAK-NEGATIVE IMPACT</td><td colspan="2" rowspan="2">POOR IMPORTANCE & HIGH FREQUENCY WEAK-POSITIVE IMPACT</td></tr>
<tr><td>2 (NOT VERY IMPORTANT)</td></tr>
<tr><td>3 (SOMEWHAT IMPORTANT)</td><td colspan="3" rowspan="3">HIGH IMPORTANCE & POOR FREQUENCY
STRONG-NEGATIVE IMPACT</td><td colspan="2" rowspan="3">HIGH IMPORTANCE & FREQUENCY
STRONG & POSITIVE IMPACT</td></tr>
<tr><td>4 (VERY IMPORTANT)</td></tr>
<tr><td>5 (EXTREMELY IMPORTANT)</td></tr>
</table>

Figures

Figure 1. Boxplots of 'frequency total' (top) and 'importance total' (bottom) scores for the NH and CI group (N=68). The boxes represent the interquartile range, i.e. the middle 50% of the observations, the whiskers are the top and bottom 25% of the scores, the horizontal bars are the medians and the circles show the outliers.

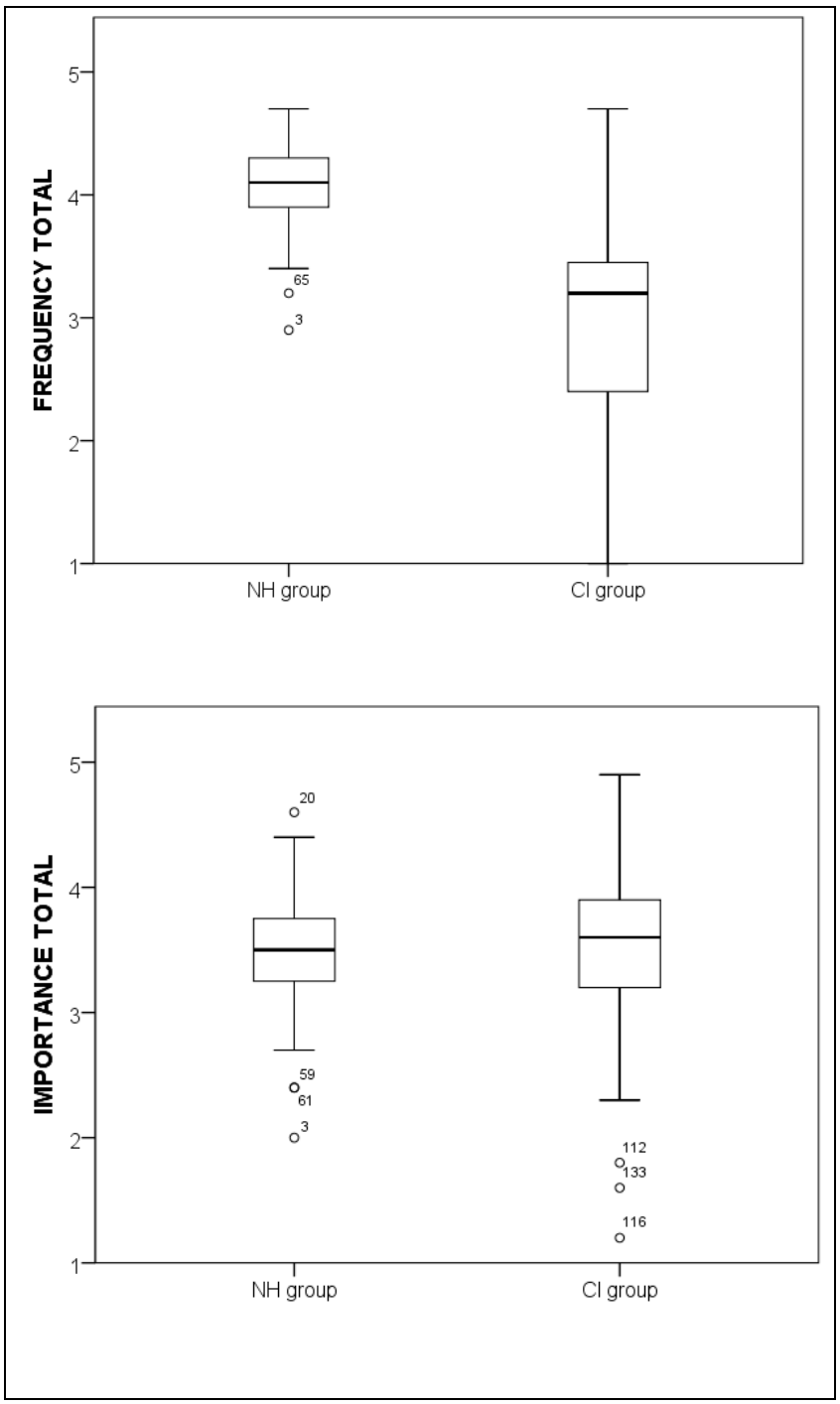
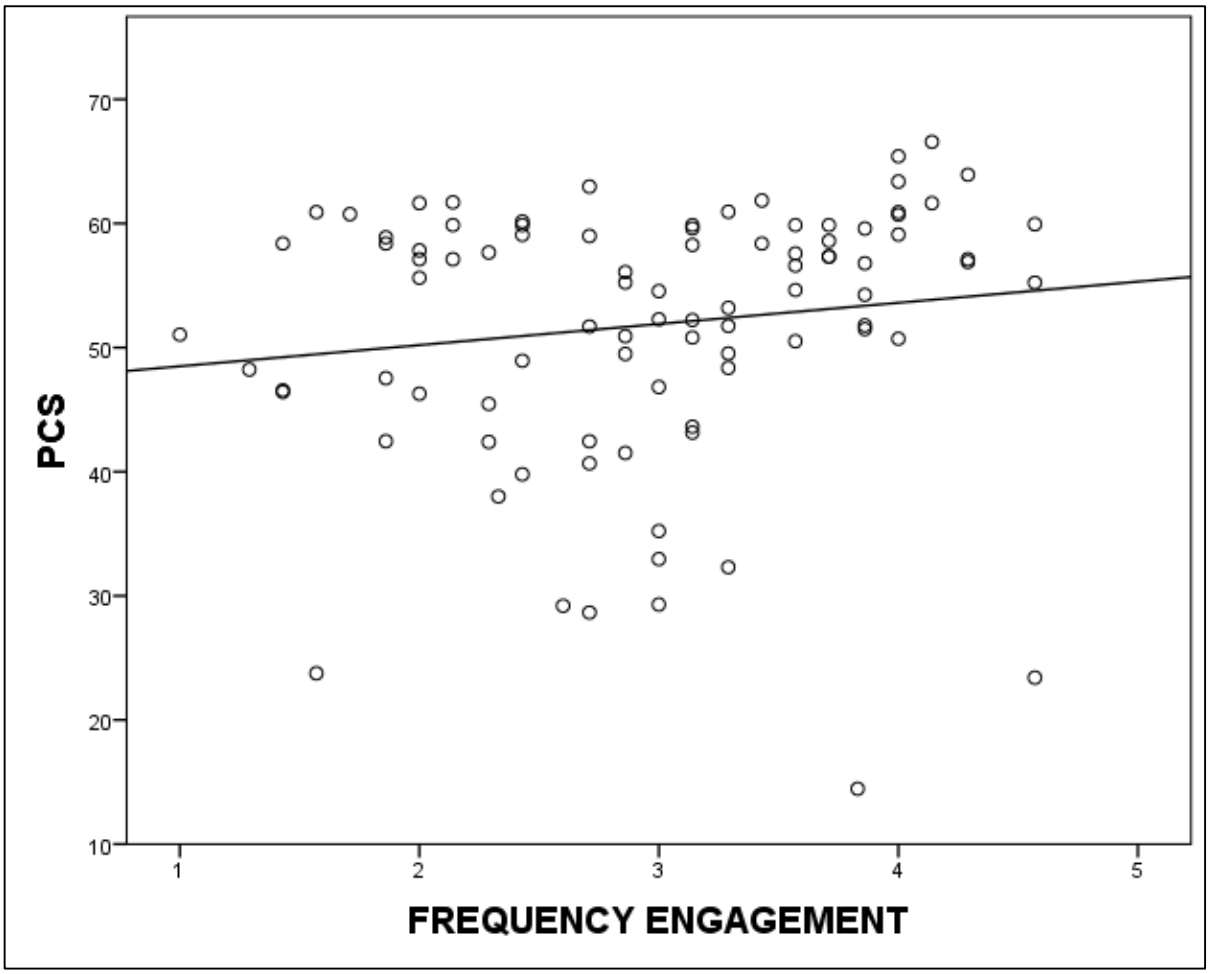


Figure 2 Scatterplot showing a significant and nearly moderate correlations between the SF12v2 PCS and the Frequency score for the MuRQoL engagement subscale (7 items) for an Importance score ≥ 3 for Engagement (N=91).



Appendix 1. The final 18 frequency and 18 importance questions of the MuRQoL questionnaire version 3 (MuRQoLv3), after item selection. For the version of the MuRQoL questionnaire available for use please see 'Data Access Statement'.

Frequency scale

	Never	Rarely	Occasionally	Frequently	Always	N/A
Music perception subscale						
1. Can you distinguish different rhythmic patterns in music?	☐	☐	☐	☐	☐	☐
2. Can you follow the melody in music (i.e. follow the melody of a song or a familiar tune)?	☐	☐	☐	☐	☐	☐
3. Can you hear differences in musical tone (i.e. how high or low music is)?	☐	☐	☐	☐	☐	☐
4. Can you recognise the words in songs?	☐	☐	☐	☐	☐	☐
5. Can you recognise the sounds of different musical instruments when they play separately ('solo')?	☐	☐	☐	☐	☐	☐
6. Can you hear the meaning of music (i.e. the emotion, why it was created or what message it is trying to get across)?	☐	☐	☐	☐	☐	☐
7. Can you hear music without effort or having to concentrate?	☐	☐	☐	☐	☐	☐
8. Can you recognise familiar music (e.g. a song, singer or tune)?	☐	☐	☐	☐	☐	☐
9. Can you judge the quality of a musical performance (e.g. singing or musical instrument playing)?	☐	☐	☐	☐	☐	☐
10. Do you feel confident that you hear music like other people do?	☐	☐	☐	☐	☐	☐
11. Does music sound in tune?	☐	☐	☐	☐	☐	☐

	Never	Rarely	Occasionally	Frequently	Always	N/A
Music engagement subscale						
12. Do you enjoy music in noisy environments when no visual cues are available (e.g. at a party, at a restaurant or in the car over the engine/road noise)?	☐	☐	☐	☐	☐	☐
13. Do you enjoy music on TV, DVD or on the computer when visual cues are available?	☐	☐	☐	☐	☐	☐
14. Do you choose to have music on in the background while doing something else (e.g. while reading, painting, doing gardening, exercising or just relaxing)?	☐	☐	☐	☐	☐	☐
15. Do you listen to music whilst travelling (e.g. in the car)?	☐	☐	☐	☐	☐	☐
16. Do you choose to listen to new music (i.e. music that you have not heard before)?	☐	☐	☐	☐	☐	☐
17. Do you attend public music events (e.g. musicals, concerts or music festivals)?	☐	☐	☐	☐	☐	☐
18. Do you sing, play a musical instrument or whistle when you are alone?	☐	☐	☐	☐	☐	☐

Importance scale

	Not important at all	Not very important	Somewhat important	Very important	Extremely important	N/A
Music perception subscale						
1. How important is it for you to be able to distinguish different rhythmic patterns in music?	☐	☐	☐	☐	☐	☐
2. How important is it for you to be able to follow the melody in music (i.e. follow the melody of a song or a familiar tune)?	☐	☐	☐	☐	☐	☐
3. How important is it for you to be able to hear differences in musical tone (i.e. how high or low music is)?	☐	☐	☐	☐	☐	☐

	Not important at all	Not very important	Somewhat important	Very important	Extremely important	N/A
4. How important is it for you to be able to recognise the words in songs?	☐	☐	☐	☐	☐	☐
5. How important is it for you to be able to recognise the sounds of different musical instruments when they play separately ('solo')?	☐	☐	☐	☐	☐	☐
6. How important is it for you to be able to hear the meaning of music (i.e. the emotion, why it was created or what message it is trying to get across)?	☐	☐	☐	☐	☐	☐
7. How important is it for you to be able to hear music without effort or without having to concentrate?	☐	☐	☐	☐	☐	☐
8. How important is it for you to be able to recognise familiar music (e.g. a song, singer or tune)?	☐	☐	☐	☐	☐	☐
9. How important is it for you to be able to judge the quality of a musical performance (e.g. singing or musical instrument playing)?	☐	☐	☐	☐	☐	☐
10. How important is it for you to feel confident that you hear music like other people do?	☐	☐	☐	☐	☐	☐
11. How important is it for you to hear music that sounds in tune?	☐	☐	☐	☐	☐	☐
Music engagement subscale						
12. How important is it for you to enjoy music in noisy environments when no visual cues are available (e.g. at a party, at a restaurant or in the car over the engine/road noise)?	☐	☐	☐	☐	☐	☐
13. How important is it for you to enjoy music on TV, DVD or on the computer when visual cues are available?	☐	☐	☐	☐	☐	☐
14. How important is it for you to have music on in the background while doing something else (e.g. while reading, painting, doing gardening, exercising or just relaxing)?	☐	☐	☐	☐	☐	☐
15. How important is it for you to listen to music whilst travelling (e.g. in the car)?	☐	☐	☐	☐	☐	☐

	Not important at all	Not very important	Somewhat important	Very important	Extremely important	N/A
16. How important is it for you to listen to new music (i.e. music that you have not heard before)?	☐	☐	☐	☐	☐	☐
17. How important is it for you to attend public music events (e.g. musicals, concerts or music festivals)?	☐	☐	☐	☐	☐	☐
18. How important is it for you to sing, play a musical instrument or whistle when you are alone?	☐	☐	☐	☐	☐	☐

Appendix 2. Items of the final MuRQoL questionnaire and recommended changes highlighted in italics. These changes are included in the version of the MuRQoL questionnaire available for use (MuRQoLv4), see 'Data Access Statement'.

Item	MuRQoL questionnaire version 3 (MuRQoLv3)	Recommended changes
1	Can you distinguish different *rhythmic patterns* in music?	Can you distinguish different *rhythms* in music?
	How important is it for you to be able to distinguish different *rhythmic patterns* in music?	How important is it for you to be able to distinguish different *rhythms* in music?
5	Can you recognise the sounds of different musical instruments *when they play separately (solo)*?	Can you recognise the sounds of different musical instruments?
	How important is it for you to be able to recognise the sounds of different musical instruments *when they play separately (solo)*?	How important is it for you to be able to recognise the sounds of different musical instruments?
13	Do you enjoy music on TV, DVD or on the computer *when visual cues are available*?	Do you enjoy music on TV, DVD or on the computer?

Item	MuRQoL questionnaire version 3 (MuRQoLv3)	Recommended changes
	How important is it for you to enjoy music on TV, DVD or on the computer *when visual cues are available*?	How important is it for you to enjoy music on TV, DVD or on the computer?
18	Do you sing, play a musical instrument or whistle *when you are alone*?	Do you sing, play a musical instrument or whistle?
	How important is it for you to sing, play a musical instrument or whistle *when you are alone*?	How important is it for you to sing, play a musical instrument or whistle?

Data access statement: The MuRQoL questionnaire (MuRQoLv4 including the recommended changes, see 'Discussion') is openly accessible from the University of Southampton repository at http://dx.doi.org/10.5258/SOTON/xxxxx under a CC-BY-NC licence. It is available free of charge for non-commercial research with acknowledgement. For additional information or instructions please contact the first author.

Acknowledgments

We are grateful to all our participants for their help. We would also like to thank Prof. Helen Cullington from the USAIS, Action on Hearing Loss, the Ear Foundation, the British Association of Music Therapy and Caroline Bartrop from the Hearing Implant Centre of Guy's and St. Thomas' Hospital (London) for their help with participants' recruitment. Funding is acknowledged from the EPSRC Doctoral Training Grant awarded to the University of Southampton, *EP/K503150/1* and by the USAIS Donations Grant.

References

Amann, E. & Anderson, I. (2014). Development and validation of a questionnaire for hearing implant users to self-assess their auditory abilities in everyday communication situations: the Hearing Implant Sound Quality Index (HISQUI19). *Acta oto-laryngologica*. 134 (9). p.pp. 915–23.

Bartel, L.R., Greenberg, S., Friesen, L.M., Ostroff, J., Bodmer, D., Shipp, D. & Chen, J.M. (2011). Qualitative case studies of five cochlear implant recipients' experience with music. *Cochlear implants international*. 12 (1). p.pp. 27–33.

van Besouw, R.M., Oliver, B.R., Grasmeder, M.L., Hodkinson, S.M. & Solheim, H. (2015).

Evaluation of an interactive music awareness program for cochlear implant recipients. *Music Perception*. p.pp. 2255479–2255483.

Bowling, A. (2014). *Research methods in health. Investigating health and health services.* 4th Ed. New York, NY: Open University Press.

Brockmeier, S.J., Fitzgerald, D., Searle, O., Fitzgerald, H., Grasmeder, M., Hilbig, S., Vermiere, K., Peterreins, M., Heydner, S. & Arnold, W. (2011). The MuSIC perception test: a novel battery for testing music perception of cochlear implant users. *Cochlear implants international*. 12 (1). p.pp. 10–20.

Brockmeier, S.J., Nopp, P., Vischer, M., Baumgartner, W., Stark, T., Schon, F., Muller, J., Braunschweig, T., Busch, R., Getto, M., Arnold, W. & Allum, D.J. (2002). Correlation of speech and music perception in postlingually deaf Combi 40/40 users. In: *Cochlear Implants: An Update*. pp. 459–464.

Calvino, M., Gavilán, J., Sánchez-Cuadrado, I., Pérez-Mora, R.M., Muñoz, E., Díez-Sebastián, J. & Lassaletta, L. (2015). Using the HISQUI29 to assess the sound quality levels of Spanish adults with unilateral cochlear implants and no contralateral hearing. *European Archives of Oto-Rhino-Laryngology*.

Chin, T. & Rickard, N.S. (2012). The Music USE (MUSE) questionnaire: an instrument to measure engagement in music. *Music Perception*. 29 (4). p.pp. 429–446.

Coelho, D.H., Hammerschlag, P.E., Bat-chava, Y. & Kohan, D. (2009). *Psychometric validity of the Cochlear Implant Function Index (CIFI): a quality of life assessment tool for adult cochlear implant users*. 10 (January). p.pp. 70–83.

Deyo, R.A., Diehr, P. & Patrick, D.L. (1991). Reproducibility and responsiveness of health status measures statistics and strategies for evaluation. *Controlled clinical trials*. 12 (4).

p.pp. S142–S158.

Drennan, W.R., Oleson, J.J., Gfeller, K., Crosson, J., Driscoll, V.D., Won, J.H., Anderson, E.S. & Rubinstein, J.T. (2014). Clinical evaluation of music perception, appraisal and experience in cochlear implant users. *International journal of audiology*. (October 2013). p.pp. 1–10.

Edwards, S.L. (2014). *Pitch perception, production and musical development of hearing impaired children*. UCL (University College London).

Field, A. (2014). *Discovering statistics using IBM SPSS Statistics*. 4th Ed. Sage Publications.

Frederigue-Lopes, N.B., Bevilacqua, M.C. & Costa, O.A. (2015). Munich Music Questionnaire: adaptation into Brazilian Portuguese and application in cochlear implant users. *CoDAS*. 27 (1). p.pp. 13–20.

Fuller, C., Mallinckrodt, L., Maat, B., Başkent, D. & Free, R. (2013). Music and quality of life in early-deafened late-implanted adult cochlear implant users. *Otology & neurotology*. 34 (6). p.pp. 1041–7.

Gfeller, K., Christ, A., Knutson, J.F., Witt, S., Murray, K.T. & Tyler, R.S. (2000). Musical Backgrounds, Listening Habits, and Aesthetic Enjoyment of Adult Cochlear Implant Recipients. *Journal of the American Academy of Audiology*. 11. p.pp. 390–406.

Gfeller, K., Driscoll, V., Smith, R.S. & Scheperle, C. (2012). The music experiences and attitutes of a first cohort of prelingually-deaf adolescents and young adults CI recipients. *Seminars in Hearing*. 33 (4). p.pp. 346–360.

Gfeller, K., Oleson, J., Knutson, J.F., Breheny, P., Driscoll, V. & Olszewski, C. (2008). Multivariate Predictors of Music Perception and Appraisal by Adult Cochlear Implant Users. *Journal of the American Academy of Audiology*. 19 (2). p.pp. 120–134.

Guyatt, G., Walter, S. & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases*. 40 (2). p.pp. 171–178.

Hinderink, J.B., Krabbe, P.F. & Van Den Broek, P. (2000). Development and application of a health-related quality-of-life instrument for adults with cochlear implants: the Nijmegen cochlear implant questionnaire. *Otolaryngology-Head and Neck Surgery*. 123 (6). p.pp. 756–65.

IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

Kang, R., Nimmons, G.L., Drennan, W., Longnion, J., Ruffin, C., Nie, K., Won, J.H., Worman, T., Yueh, B. & Rubinstein, J. (2009). Development and validation of the University of Washington Clinical Assessment of Music Perception test. *Ear and hearing*. 30 (4). p.pp. 411–8.

Landis, J.R. & Koch, G.G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*. 33. p.pp. 159–174.

Lassaletta, L., Castro, A., Bastarrica, M., Pérez-Mora, R., Herrán, B., Sanz, L., de Sarriá, M.J. & Gavilán, J. (2008). Musical perception and enjoyment in post-lingual patients with cochlear implants. *Acta otorrinolaringológica española*. 59 (5). p.pp. 228–34.

Limb, C.J. & Roy, A.T. (2013). Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hearing research*. p.pp. 1–14.

Loeffler, C., Aschendorff, A., Burger, T., Kroeger, S., Laszig, R. & Arndt, S. (2010). Quality of Life Measurements after Cochlear Implantation. *The Open Otorhinolaryngology Journal*. 4. p.pp. 47–54.

Looi, V., Gfeller, K. & Driscoll, V. (2012). Music appreciation and training for cochlear implant recipients: A review. *Seminars in Hearing*. 33 (4). p.pp. 307–334.

Looi, V., McDermott, H., McKay, C. & Hickson, L. (2008). Music perception of cochlear implant users compared with that of hearing aid users. *Ear and hearing*. 29 (3). p.pp. 421–34.

Looi, V. & She, J. (2010). Music perception of cochlear implant users: a questionnaire, and its implications for a music training program. *International journal of audiology*. 49 (2). p.pp. 116–28.

Maarefvand, M., Marozeau, J. & Blamey, P.J. (2013). A cochlear implant user with exceptional musical hearing ability. *International journal of audiology*. 52 (6). p.pp. 424–32.

McHorney, C. a & Tarlov, a R. (1995). Individual-patient monitoring in clinica practice: are available health status survey adequate? *Quality of Life Research*. 4 (93). p.pp. 293–307.

Migirov, L., Kronenberg, J. & Henkin, Y. (2009). Self-reported listening habits and enjoyment of music among adult cochlear implant recipients. *The Annals of otology, rhinology, and laryngology*. 118 (5). p.pp. 350–5.

Nunnally, J.C. & Bernstein, I.H. (1994). *Psychometric theory*. Third. McGraw-Hill.

Renwick, R., Nourhaghighi, N., Manns, P.J. & Rudman, D.L. (2003). Quality of life for people with physical disabilities: a new instrument. *International journal of rehabilitation research.* 26 (4). p.pp. 279–87.

Revicki, D., Hays, R.D., Cella, D. & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology*. 61 (2). p.pp. 102–109.

Schäfer, T., Sedlmeier, P., Städtler, C. & Huron, D. (2013). The psychological functions of music listening. *Frontiers in psychology*. 4 (August). p.p. 511.

Streiner, D.L., Norman, G.R. & Cairney, J. (2015). *Health Measurement Scales. A practical guide to their development and use.* 5th Ed. Oxford University Press.

Veekmans, K., Ressel, L., Mueller, J., Vischer, M. & Brockmeier, S.J. (2009). Comparison of music perception in bilateral and unilateral cochlear implant users and normal-hearing subjects. *Audiology & neuro-otology*. 14 (5). p.pp. 315–26.

Ware, J.E., Kosinski, M. & Keller, S.D. (1996). A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. *Medical Care*. 34 (3). p.pp. 220–233.

WHOQOL (1993). Study protocol for the World Health Organization project to develop a Quality of Life assessment instrument (WHOQOL). *Quality of life research*. 2 (2). p.pp. 153–9.

Wright, R. & Uchanski, R.M. (2012). Music Perception and Appraisal: Cochlear Implant Users and Simulated CI Listening. *Journal of American Academy of Audiology*. 23 (5). p.pp. 350–379.

Yang, Y., Longworth, L. & Brazier, J. (2013). An assessment of validity and responsiveness of generic measures of health-related quality of life in hearing impairment. *Quality of life research*.

Zhao, F., Bai, Z. & Stephens, D. (2008). The relationship between changes in self-rated quality of life after cochlear implantation and changes in individual complaints. *Clinical otolaryngology*. 33 (5). p.pp. 427–34.