International Psychogeriatrics (2014), 26:6, 1011–1019 © International Psychogeriatric Association 2014
doi:10.1017/S1041610214000180

A preliminary psychometric evaluation of Music in Dementia Assessment Scales (MiDAS)

Orii McDermott,[1,2,3] Vasiliki Orgeta,[1] Hanne Mette Ridder[2] and Martin Orrell[1,4]

[1] *Division of Psychiatry, Faculty of Brain Sciences, University College London, UK*
[2] *Doctoral Programme in Music Therapy, Institute for Communication and Psychology, Aalborg University, Denmark*
[3] *Central and North West London NHS Foundation Trust, London, UK*
[4] *Research and Development, North East London NHS Foundation Trust, London, UK*

ABSTRACT

Background: Music in Dementia Assessment Scales (MiDAS), an observational outcome measure for music therapy with people with moderate to severe dementia, was developed from qualitative data of focus groups and interviews. Expert and peer consultations were conducted at each stage of the scale development to maximize its content validity. This study aimed to evaluate the psychometric properties of MiDAS.

Methods: Care home residents with dementia attended weekly group music therapy for up to ten sessions. Music therapists and care home staff were requested to complete weekly MiDAS ratings. The Quality of Life Scale (QoL-AD) was completed at three time-points.

Results: A total of 629 (staff = 306, therapist = 323) MiDAS forms were completed. The statistical analysis revealed that MiDAS has high therapist inter-rater reliability, low staff inter-rater reliability, adequate staff test-retest reliability, adequate concurrent validity, and good construct validity. High factor loadings between the five MiDAS Visual Analogue Scale (VAS) items, levels of *Interest*, *Response*, *Initiation*, *Involvement*, and *Enjoyment*, were found.

Conclusions: This study indicates that MiDAS has good psychometric properties despite the small sample size. Future research with a larger sample size could provide a more in-depth psychometric evaluation, including further exploration of the underlying factors. MiDAS provides a measure of engagement with musical experience and offers insight into who is likely to benefit on other outcomes such as quality of life or reduction in psychiatric symptoms.

Key words: music therapy, dementia, outcome measure, psychometric evaluation, assessment

Introduction

Music-based interventions are popular in clinical practice with people with dementia. Systematic reviews on psychosocial treatments of behavior and psychological symptoms in dementia have found sufficient evidence to support the effectiveness of individually tailored music interventions (O'Connor *et al.*, 2009a; 2009b). A number of systematic reviews on music therapy have also found consistent evidence for short-term reductions in behavioral disturbance and improved mood in people with dementia (e.g. Koger *et al.*, 1999; Vink *et al.*, 2013; Ueda *et al.*, 2013). On the other hand, music therapy studies have been criticized for the lack of evidence of long-term benefits (Livingston *et al.*, 2005) and for the weaknesses in study designs (McDermott *et al.*, 2013; Vink *et al.*, 2013). All the reviews agree that the potential benefits of music-based interventions for the well-being of people with dementia are promising, but the effects of the interventions require further investigations.

There is also concern whether the outcome measures used in the current music therapy studies are most appropriate to evaluate the intervention. Our recent systematic review (McDermott *et al.*, 2013) found that the Neuropsychiatric Inventory (NPI; Cummings *et al.*, 1994), the Cohen-Mansfield Agitation Inventory (CMAI; Cohen-Mansfield, 1986), and various depression scales are most frequently used to evaluate the impact of music therapy on behavioral and psychological symptoms. On the other hand, there is consistent anecdotal clinical evidence that people with dementia benefit from "positive experiences" of music therapy, for instance, opportunities for

Correspondence should be addressed to: Orii McDermott, Division of Psychiatry, Faculty of Brain Sciences, University College London, Charles Bell House, 67–73 Riding House Street, London W1W 7EJ, UK. Phone: +44-(0)20-7679-9064; Fax: +44-(0)20-7679-9426. Email: o.mcdermott@ucl.ac.uk. Received 4 Sep 2013; revision requested 7 Dec 2013; revised version received 20 Jan 2014; accepted 21 Jan 2014. First published online 13 February 2014.

self-expression, increased self-awareness, and social interactions with fellow participants. This raises a question if music therapy should only be seen as a "treatment" to fix neuropsychiatric symptoms.

Music therapy is particularly disadvantaged by the lack of client-specific, clinically appropriate, and psychometrically evaluated outcome measures. York (1994) developed the Residual Music Skills Test for people with probable Alzheimer's disease (AD). A test-retest showed a high correlation ($r = 0.916$, $p < 0.000$) (York, 2000). However, testing the musical skills of clients may not be the most appropriate outcome measure to evaluate the intervention. Similarly, an assessment tool to evaluate behavioral responses of people while listening to "taped music" (Glynn, 1992) is not the most appropriate measure to evaluate the benefits of music therapy for people with dementia, even though the study reported a high inter-rater reliability ($r = 0.970$–0.999). In some studies (e.g. Ashida, 2000; Brotons and Marti, 2003) researchers used their own scales, but these were not rigorously developed or evaluated for reliability and validity.

Music in Dementia Assessment Scales (MiDAS) (McDermott *et al.*, unpublished work), an observational outcome measure for music therapy with people with moderate to severe dementia, was developed from qualitative data of focus groups and interviews conducted with people with dementia, family carers, care home staff, and music therapists. Explicit qualitative data analysis methods utilizing the general inductive approach (Thomas, 2006) and the long-table approach (Kruger and Casey, 2000) ensured the rigor and transparency of the scale development. Expert and peer consultations were conducted at each stage of the scale development to maximize the content validity of MiDAS. The pilot MiDAS was field-tested by a music therapist and care home staff for six weeks. Feedback from the therapist and staff and further comments from external music therapists were collated, and the revised MiDAS form was produced.

The use of qualitative data for the MiDAS development was to ensure its clinical relevance. The aim of this study was to evaluate the psychometric properties of MiDAS.

Methods

Design

Care home residents with moderate to severe dementia attended weekly group music therapy. Music therapists and care home staff were requested to complete weekly MiDAS ratings. Quality of Life in Alzheimer's disease (QoL-AD; Logsdon *et al.*, 1999) was completed at baseline, mid-treatment, and end-treatment by care home staff. Statistical analysis was conducted with SPSS version 21. Inter-rater reliability, test-retest reliability, internal consistency, concurrent validity, and construct validity of MiDAS were evaluated.

Table 1. Examples of behavior and mood for VAS items provided on the MiDAS form

Interest	Did he/she show his/her interest in an activity or other people around him/her?
	Did his/her posture or facial expression change if activities or music catch his/her attention?
	Did he/she become animated if activities or music catch his/her attention?
Response	Did his/her facial expression or body movements indicate his/her awareness of staff or therapist?
	Did he/she make eye contact with staff, therapist, or other group members?
	Did he/she join in conversation, music making, or make vocal sound?
Initiation	Did he/she try to communicate with staff, therapist, or other group members?
	Did he/she start conversation, start music making, or initiate vocalization?
	Did he/she talk about his/her life experiences (reminiscence), or mention music meaningful to them?
Involvement	Did he/she become engaged in conversation, music making, or any forms of communication?
	Did he/she show his/her enthusiasm in activities that interest him/her?
Enjoyment	Smiling, laughing, brighter mood
	Playfulness, sense of humor
	Relaxed mood

Music in Dementia Assessment Scales (MiDAS)

The MiDAS consists of the following five Visual Analogue Scale (VAS) items: *Interest*, *Response*, *Initiation*, *Involvement*, and *Enjoyment*. Each VAS comprises a 100-mm line without intervals, with the two extremes of the scale labeled as "None at all" and "Highest." Examples of behavior and mood of the person that a rater should look for before scoring each VAS are provided on the MiDAS form (Table 1). The MiDAS form also includes a supplementary checklist of six notable reactions from the person (*agitation/aggression*, *withdrawn/low in mood*, *restless/anxious*, *relaxed mood*, *attentive/interested*, and *cheerful/smiling*) and a space for rater's comments to provide further information when necessary. The checklist of other major reactions should be used

to aid clinical interpretations of the VAS items when MiDAS is used for routine clinical evaluation or in a qualitative study. The checklist items are independent of the MiDAS scores. The maximum score for each VAS is 100, thus the maximum MiDAS score per rating is 500.

MiDAS rating procedure

Music therapists and care home staff were requested to complete weekly MiDAS ratings. Before a resident attended his/her music therapy, a staff member with sufficient knowledge of the resident was asked to complete a *Before* rating based on the average presentation of the resident on the day. An *After* form was completed by the same staff member a few hours after the resident's music therapy on the same day before staff finished their shift work. The knowledge about individual residents and the experiences of working in care homes A and B varied greatly among staff members. Permanent staff members, particularly registered nurses and healthcare assistants who provided regular daily care to the residents were sought as MiDAS raters whenever possible. Music therapists were asked to complete a *Beginning* form and a *During* form immediately after each session. The *Beginning* rating was based on the therapist's observation of the resident during the first five minutes of the session, while *During* rating was based on the observation of the most clinically significant five minutes ("the best five minutes") of that session. It was left to the therapist's clinical judgment to determine as to what should be considered as "the most clinically significant five minutes" of that session, since this would vary depending on each participant's clinical goals. For some this may mean an increased awareness of the therapist and a longer two-way musical interaction. For others, the "best five minutes" may imply willingness to collaborate with fellow group members and share their ideas. Thus, four MiDAS ratings were to be completed per resident per session to assess possible changes in their mood and behavior following music therapy.

Study sample

Residents who met Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV; American Psychiatric Association, 1994) criteria for dementia in two National Health Service (NHS) care homes (care homes A and B) were recruited for weekly music therapy group for up to ten sessions. Residents with dementia who also had other psychiatric conditions (e.g. schizophrenia) were included. The music therapy group comprised four to six residents, and each session was guided by the clinical judgments of two music therapists, Rosslyn Bender (RB) and Maria Radoje (MR), having over ten years of experience with this client group. Both therapists used an active music therapy model where clients are encouraged to explore musical instruments and/or their voice freely. The focus of the sessions included encouragement of musical and verbal interactions between the group members, as well as supporting the development of self-awareness of an individual. The use of pre-composed songs combined with exploratory improvisation was frequently used. Ethics approval from the National Research Ethics Service (NRES Committee London East, REC reference 11/LO/0596) was granted in June 2011.

Reliability

Inter-rater reliability examines the degree of objectivity by correlating two sets of ratings by two raters completed at any given time. To assess therapist's inter-rater reliability, Orii McDermott (OM) attended RB's music therapy group in care home B as a participant observer for six weeks. OM and RB completed therapist MiDAS ratings independently immediately after each session. For staff inter-rater reliability, an associate health practitioner and an activity coordinator in care home A were assigned as staff inter-raters by the unit manager. They were encouraged to rate the residents independently on the day of their music therapy as regularly as their work patterns permitted.

Test-retest reliability measures the repeatability of the test and involves administrating the test twice to the same population group. Staff test-retest reliability was evaluated through permanent members of care home B staff completing *Before* ratings twice on the same residents one week apart on eight separate occasions.

Intraclass Correlation Coefficients (ICC) Case 2 two-way random effects model with absolute agreement (McGraw and Wong, 1996) was utilized, and single scores were quoted for inter-rater reliability and test-retest reliability results. This model was chosen on the basis that the use of two-way random effects ANOVA and single scores allowed random participants being assessed by random judges at any single occasion, thus maximizing the generalizability of MiDAS.

Internal consistency measures the degree of correlations between the items on the same test, and was examined using Cronbach's α.

Validity

Face and content validity has already been established during the MiDAS development through the use of consensus methods, expert and peer consultations, and by collating feedback from

therapists and care home staff during the refinement stage of the pilot MiDAS.

To evaluate concurrent validity of MiDAS, the QoL-AD (Logsdon *et al.*, 1999) was administrated at three time-points (base-line, mid-treatment, end-treatment). QoL-AD, Alzheimer's disease-related quality of life is a Likert scale consisting of 13 items, including physical health, mood, family, ability to do chores around the house, and life as a whole. QoL-AD can be administrated as a self-rating scale or a proxy measure by a carer. QoL-AD was chosen for three reasons. First, there was no "gold-standard" music therapy outcome measure to compare with MiDAS, and the nature of QoL measures that assess the overall well-being of a person was closer to the underlying concept of MiDAS than the nature of the measures that assess specific neuropsychiatric symptoms. Second, QoL-AD has been utilized successfully in previous studies with people with moderate to severe dementia (e.g. Hoe *et al.*, 2005; Spector and Orrell, 2006). Third, QoL-AD was recommended in a European consensus on outcome measures for psychosocial intervention research in dementia care (Moniz-Cook *et al.*, 2008). Spearman's correlation coefficient was chosen to examine concurrent validity.

Construct validity involves measurement of psychological attributes (Nunnally and Berstein, 1994), including an investigation of the empirical relationships between the measures and an interpretation of the empirical evidence of how it clarifies the construct validity of a particular measure (Carmines and Zeller, 1969). Construct validity of MiDAS was evaluated using exploratory factor analysis and by examining inter-relationships between the MiDAS items. Although it is widely accepted that a larger sample size would provide more trustworthy results for factor analysis, there is a lack of agreement on a recommended sample size. Hogarty *et al.* (2005) argued that other elements, such as participants-to-variables ratio and communality of variance, should also be taken into consideration. As this study had a small sample size, squared multiple correlation (SMC) for each variable to estimate communality values, as well as Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy and the Bartlett's test of sphericity (Bartlett, 1954) were evaluated for the appropriateness of factor analysis before conducting Principal Component Analysis (PCA).

Statistical analysis

Statistical analysis was conducted with SPSS version 21. Due to the data characteristics, square root transformation of the original data was conducted to evaluate ICC for inter-rater reliability and test-retest reliability. Tests on non-transformed data (original data) were also conducted to assess if there were any major differences between the results. For the same reason, both Principal Axis Factoring (PAF), a descriptive method suitable when the assumption of normality cannot be established, and PCA were conducted for the evaluation of concurrent validity to compare the results. Because of the small sample size, repeated observations of the same residents at different times were treated as unique observations when evaluating inter-rater reliability, concurrent validity, and construct validity (factor analysis test 1). This method was discussed with two independent statisticians and deemed suitable for the purpose of this study. The second factor analysis was conducted on the mean MiDAS scores of each participant (factor analysis test 2). Before conducting test 2, squared multiple correlation (R^2) was used as SMC to determine the communality and appropriateness for factor analysis, as suggested by Field (2009).

Results

A total of 19 residents (5 males, and 14 females) in the two care homes attended weekly group music therapy for up to 10 sessions (mean attendance 7 sessions). The residents were White British/White European (17) or Black British (2) with a mean age of 81 years (SD 7.7). They were either single (6), married (3), separated/divorced (3), or widowed (7). All had diagnosis of dementia: Alzheimer's disease (10), vascular dementia (5), and other dementias (4). The mean stay in the care homes was 20.1 months (SD 14.2).

A total of 629 (staff = 306, therapist = 323) MiDAS forms were completed and included in the psychometric evaluation of MiDAS. The 306 MiDAS forms were completed by a total of 31 nursing staff: consisting of qualified nurses, healthcare assistants, and an activity coordinator. The mean of each five VAS item was: *Interest* 50.17 (SD = 26.96), *Response* 51.57 (SD = 28.32), *Initiation* 44.34 (SD = 30.10), *Involvement* 46.63 (SD = 29.58), and *Enjoyment* 46.17 (SD = 29.95). The mean total score was 238.87 (SD = 136.38).

Reliability

Inter-rater reliability (therapist = 52 paired forms, staff = 27) and test-retest reliability (staff = 31): Due to large SD, square root transformation, rather than log transformation, was chosen because all the VAS items (*Interest*, *Response*, *Initiation*, *Involvement* and *Enjoyment*) had scores of 0 (13/629 MiDAS forms, 11/629, 38/629, 18/629, 15/629 respectively), and these scores would have been

Table 2. Therapist and staff inter-rater reliability and test-retest reliability

	INTER-RATER (THERAPIST) ICC (N = 52) (CI)	INTER-RATER (STAFF) ICC (N = 27) (CI)	TEST-RETEST (STAFF) ICC (N = 31) (CI)
Interest	0.791* (0.656, 0.876)	0.127 (−0.275, 0.484)	0.571* (0.288, 0.766)
Response	0.789* (0.659, 0.873)	0.189 (−0.215, 0.531)	0.498** (0.181, 0.721)
Initiation	0.768* (0.626, 0.860)	0.351** (−0.037, 0.643)	0.665* (0.414, 0.822)
Involvement	0.820* (0.707, 0.893)	0.362** (−0.023, 0.651)	0.614* (0.333, 0.794)
Enjoyment	0.787* (0.643, 0.875)	0.205 (−0.185, 0.538)	0.609* (0.327, 0.791)

Notes: ICC Case 2 model: two-way random ANOVA, absolute agreement, single measures, 95% confidence interval (CI).
*$p < 0.001$, **$p < 0.05$.

excluded from analysis if log transformation was utilized. Skewness and kurtosis of the five VAS items were normalized with square root transformation. Since the differences between the results of the transformed data and the original data were not significant (ICC differences: range –0.024 to 0.003, mean: –0.008, SD: 0.011 for therapist inter-rater reliability), the original scores were retained. The results of inter-rater reliability and test-retest reliability are shown on Table 2.

Inter-rater reliability (therapist): A total of ten residents attended during therapist inter-rater sessions: seven attended all the three sessions that OM observed, two residents attended two out of three sessions, and one attend one of the three sessions. In total, 52 paired forms for ten residents were completed. ICC (two-way random model with absolute agreement) revealed high correlations (range: 0.768–0.820).

Inter-rater reliability (staff) and test-retest reliability: Twenty-seven paired forms were completed by two staff inter-raters at care home A. Very low ICC was found: ranging from 0.127 to 0.362. In order to determine whether the low score was due to the low reliability of MiDAS itself, or whether due to the difference in staff roles, we conducted test-retest on staff *Before* forms that were completed by any staff member over any consecutive weeks. Fifteen paired-forms were found among the 290 forms completed by staff. In addition, OM asked permanent staff members in care home B to complete further MiDAS test-retest (one week apart) on selected residents they knew well on eight separate occasions. In total, 31 paired forms were examined. Acceptable correlations were found, ICC ranged from 0.498 to 0.609.

Internal consistency

Cronbach's α showed high correlations ($\alpha = 0.967$) between the five VAS items (n = 629).

Concurrent validity (n = 36)

QoL-AD was administrated at baseline, mid-treatment, and end-treatment. The three results were combined (n = 36) to evaluate an overall correlation, resulting in Spearman's $\rho = 0.480$ ($p = 0.003$). Due to missing data, analysis was conducted on 10 residents (baseline), 14 residents (mid-treatment), and 12 residents (end-treatment) out of the 19 residents who participated in music therapy sessions. The outcome of Spearman's rank correlation coefficient for each time-point was acceptable: 0.524 (base-line), 0.469 (mid-treatment), and 0.474. (end-treatment).

Construct validity

Factor analysis (Test 1: 629 MiDAS ratings)

Individual MiDAS ratings were treated as unique observations and all the 629 forms were included for factor analysis test 1. The KMO measure of sampling adequacy and the Bartlett's test of sphericity (Bartlett, 1954) were carried out. The KMO value was 0.895, which exceeds the value of 0.6 suggested by Pallant (2005). The Bartlett's test of sphericity was statistically significant ($p < 0.001$), supporting the factorability of the correlation matrix for PCA. A PAF was also conducted. The outcomes of KMO, Barlett's Test, and the factor loadings were identical for PCA and PAF. Eigenvalues greater than 1 were used to select factors. Five components were extracted for the initial eigenvalues but only one was retained for both PCA and PAF, with a single underlying factor, which explained 88.5% of the variance. One component was interpreted as engagement with music. Correlation between *Response* and *Involvement* was the highest (0.921), and the lowest correlation was observed for *Initiation* and *Enjoyment* was 0.791 (Table 3).

Factor analysis (Test 2: 19 residents)

R^2-scores for the five VAS items ranged from 0.568 to 0.946 (mean = 0.793), exceeding the recommended mean level of communality above 0.7 (MacCallum *et al.*, 1999). The KMO value

Table 3. Factor loadings (Principal Component Analysis) test 1 (n = 629)

	INTEREST	RESPONSE	INITIATION	INVOLVEMENT	ENJOYMENT
Interest	1.00				
Response	0.921*	1.00			
Initiation	0.821*	0.857*	1.00		
Involvement	0.859*	0.888*	0.891*	1.00	
Enjoyment	0.843*	0.848*	0.791*	0.845*	1.00

Notes: Identical results for PCA and PAF.
*$p < 0.001$.

Table 4. Factor loadings (Principal Component Analysis) test 2 (n = 19)

	INTEREST	RESPONSE	INITIATION	INVOLVEMENT	ENJOYMENT
Interest	1.00				
Response	0.947*	1.00			
Initiation	0.869*	0.928*	1.00		
Involvement	0.901*	0.955*	0.973*	1.00	
Enjoyment	0.844*	0.885*	0.754*	0.828*	1.00

Note: *$p < 0.001$.

was 0.824, and the Bartlett's test of Sphericity was statistically significant ($p < 0.001$). These results suggested that factor analysis was acceptable for this study despite its small sample size. Five components were extracted for the initial eigenvalues but only one was retained, with a single underlying factor, which explained 91.1% of the variance. The pattern of the correlations was very similar to that of test 1, but higher correlations between the items were observed in all but two of the paired correlations (*Initiation* and *Enjoyment*, and *Involvement* and *Enjoyment*) (Table 4).

Discussion

This study evaluated the psychometric properties of MiDAS. MiDAS is unique in the sense that it is an observational, dementia-specific VAS developed to measure musical experiences of a person with dementia. Rigorous development of MiDAS using consensus methods, peer and expert consultations ensured its content validity. Staff and therapists who participated in the main study confirmed the clinical relevance of MiDAS, and the preliminary psychometric evaluation revealed that MiDAS has acceptable to good psychometric properties. At the same time, some aspects of the results may need to be interpreted cautiously.

Methodological problems

Although a large number of MiDAS forms were completed for this study, the statistical analysis method was affected by the small sample size (n = 19) and the missing data (e.g. uncompleted QoL-AD forms). RB and OM completed 52 MiDAS forms each for the purpose of therapist inter-rater reliability. However, just seven out of ten group members attended all three sessions that OM observed, resulting in only seven patients eligible for ICC analysis. Since the main purpose of the inter-rater reliability was to assess the reliability of MiDAS as an instrument, and not to evaluate clinical changes in the participants, it was deemed acceptable to treat the 52 paired MiDAS forms as individual cases for this study, but further evaluation is required to confirm the results. Two sets of factor analysis were conducted to accommodate the limitation of using repeated observations of the same residents as independent observations. However, there is a possibility that the small sample size had an impact on the reduced observable variance, affecting the initial eigenvalues and the number of components extracted. Further evaluation with a larger sample size is required. Finally, all the QoL-AD forms were completed by staff, as the residents could not complete the self-rating forms. Concurrent validity of MiDAS should perhaps be tested further with people with mild to moderate dementia who could complete the self-rating QoL-AD.

Challenges of using MiDAS in clinical settings

The reason for the low staff inter-rater reliability may be partially due to the differences in the professional roles of the two raters (an associate health practitioner and an activity coordinator). The health practitioner worked on shifts and mainly

worked with the residents on an individual basis, while the activity coordinator provided therapeutic group activities. The health practitioner explained that she spoke to the residents before completing MiDAS forms if they appeared asleep or had been showing little response to external stimuli on the day. The activity coordinator acknowledged the lack of time spent with individual residents to assess subtle awareness or interest from the residents. This may explain the higher inter-rater reliability for *Initiation* and *Involvement* that residents' reactions were possibly more visible, while evaluating *Interest* and *Response* of less active residents required more staff time and possibly a deeper knowledge of each individual resident. Nurses and healthcare assistants who help with the residents' daily care were considered as the key staff MiDAS raters, but involving permanent staff members with in-depth knowledge of the residents was sometimes not possible due to their job responsibilities and time constraints.

Another major challenge of using MiDAS relates to the use of VAS as an observational outcome measure. MiDAS is a relative scale; for instance, one resident's highest level of interest will differ from that of another resident, and the person's optimum levels will shift with progress in dementia. Judging each resident's optimum levels of function and scoring MiDAS relies on a rater's observational skills, his/her relationship with the resident, and their willingness to stop and reflect. It was particularly difficult to have consistent staff raters in care home B. The fact that some residents were rated by four or five different staff members over ten weeks may have had some impact on the consistency of their scores. RB expressed her concern regarding the validity of staff scores after observing how some staff members were reluctant to complete the forms as they saw this process as an additional task they were asked to complete. This highlights the importance of training new MiDAS raters to clarify the purpose and procedure of ratings.

Limitations of MiDAS

Factor loadings revealed high correlations between the five VAS items. The high internal consistency was reflected in Cronbach's α (0.967). These results may have been influenced by the small number of items ($n = 5$), and may also be an indication that the five VAS items are too similar to discriminate their characteristics from each other. It is also possible that the understanding of each of the items varied greatly between the raters, resulting in averaged mid-range scores for all items, and hence high correlations. Since the main aim of MiDAS is to measure what people with dementia value in music, it can be argued having one construct (engagement with music) and using the total score of the five VAS items as the main outcome does not devalue the clinical relevance of MiDAS. Nevertheless, each VAS item was chosen carefully following the rigorous qualitative analysis. Further exploration of the underlying concepts for each VAS item through peer and expert consultations may be useful.

Future recommendations

Having different staff raters every week reduces the reliability of the changes in each resident's MiDAS scores over time. Nevertheless, this will be the reality in most care homes where shift working is the norm. The psychometric properties of MiDAS need to be evaluated further with a randomized controlled study with a larger sample size in several different care settings where the ethos of care and the levels of staff training are different.

It may be relevant to develop and evaluate a self-rating MiDAS. Likewise, it may also be feasible to develop a MiDAS family carer version that can be used in a music intervention study for people with dementia living in the community. The comparison of self-rating MiDAS scores with proxy MiDAS scores by family carers may clarify what observable clinical changes are and highlight the differences between internal musical experiences for people with dementia and perception of observers. Both the family carer and self-rating versions should include the views of people with dementia and their families living in the community.

Conclusions

MiDAS was rigorously developed using qualitative methods and consensus approaches. This study indicates that MiDAS has adequate psychometric properties on a range of attributes even though the sample size was small. This preliminary evaluation revealed that MiDAS has high therapist inter-rater reliability, low staff inter-rater reliability, adequate staff test-retest reliability, adequate concurrent validity, and good construct validity. The high correlations between the five VAS items highlighted the need for ensuring the clarification of the differences between the items particularly with new MiDAS raters, but it also suggests that adding together the MiDAS items scores produces a meaningful summary score. Future research with a larger sample size could provide a more in-depth psychometric evaluation, including further exploration of potential underlying factors. The majority of music therapy studies in dementia focus on the reduction of neuropsychiatric symptoms and not on the increased well-being of people

with dementia. MiDAS challenges this culture by aiming to capture some of the aspects of musical experiences that people with dementia themselves would say are meaningful to measure. Much of the research on music in dementia is inconclusive and has failed to offer an explanation of who may or may not benefit. MiDAS provides a measure of engagement with the musical experience and therefore may provide a unique insight into who is likely to benefit in terms of improvement on other outcomes such as quality of life or reduction in psychiatric symptoms.

Conflict of interest

None.

Description of authors' roles

O. McDermott was the main researcher responsible for designing and conducting the study, data collection, data analysis, and writing the paper. V. Orgeta was involved in data analysis and provided statistical support. H.M. Ridder was involved in the development of MiDAS and provided clinical expert opinion. M. Orrell supervised the designing and conducting of the study and provided expert advise during data analysis and writing the paper. All authors have approved the final draft of this paper.

Acknowledgments

We would like to thank music therapists Rosslyn Bender and Maria Radoje, and all the care home staff for their time and contributions for the study. We would also like to express our appreciation to the residents in the two care homes and for the support of Central and North West London NHS Foundation Trust.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders*, 4th edn. Washington: American Psychiatric Association.

Ashida, S. (2000). The effect of reminiscence music therapy sessions on changes in depressive symptoms in elderly persons with dementia. *Journal of Music Therapy*, 37, 170–182.

Bartlett, M. S. (1954). A note on multiplying factors for various chi-squared approximations. *Journal of the Royal Statistical Society, Series B* 16, 296–298.

Brotons, M. and Marti, P. (2003). Music therapy with Alzheimer's patients and their family caregivers: a pilot project. *Journal of Music Therapy*, 40, 138–150.

Carmines, E. and Zeller, R. (1969). *Reliability and Validity Assessment*. Sage University Papers. Thousand Oaks CA: Sage.

Cohen-Mansfield, J. (1986). Agitated behaviors in the elderly. II. Preliminary results in the cognitively deteriorated. *Journal of the American Geriatrics Society*, 34, 722–727.

Cummings, J., Mega, M., Gray, K., Rosenberg-Thompson, S., Carusi, D. and Gornbein, J. (1994). The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology*, 44, 2308–2314. doi:10.1212/WNL.44.12.2308

Field, A. (2009). Exploratory factor analysis. In *Discovering Statistics Using SPSS*, 3rd edn. (pp.627–685). London: Sage.

Glynn, N. J. (1992). The music therapy assessment tool in Alzheimer's patients. *Journal of Gerontological Nursing*, 18, 3–9.

Hoe, J., Katona, C., Roch, B. and Livingston, G. (2005). Use of the QOL-AD for measuring quality of life in people with severe dementia – the LASER-AD study. *Age and Ageing*, 34, 130–135. doi:10.1093/ageing/afi030.

Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M. and Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: the influence of sample size, communality and overdetermination. *Educational and Psychological Measurement*, 65, 202–226.

Koger, S. M., Chapin, K. and Brotons, M. (1999). Is music therapy an effective intervention for dementia? A meta-analytic review of literature. *Journal of Music Therapy*, 36, 2–15.

Kruger, R. and Casey, M. (2000). *Focus Groups. A Practical Guide for Applied Research*, 3rd edn. Thousand Oaks, CA: Sage.

Livingston, G., Johnston, K., Katona, C., Paton, J. and Lyketsos, C. (2005). Systematic review of psychological approaches to the management of neuropsychiatric symptoms of dementia. *American Journal of Psychiatry*, 162, 1996–2021. doi:10.1176/appi.ajp.162.11.1996.

Logsdon, R., Gibbons, L., McCurry, S. and Teri, L. (1999). Quality of life in Alzheimer's disease: patient and caregiver reports. *Journal of Mental Health and Aging*, 5, 21–32.

MacCallum, R. C., Widaman, K. F., Zhang, S. and Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4, 84–99.

McDermott, O., Crellin, N., Ridder, H. M. and Orrell, M. (2013). Music therapy in dementia: a narrative synthesis systematic review. *International Journal of Geriatric Psychiatry*, 28, 781–794. doi:10.1002/gps.3895.

McDermott, O., Orrell, M. and Ridder, H. M. (unpublished work). The importance of music for people with dementia: the perspectives of people with dementia, family carers, staff and music therapists. *Aging & Mental Health*. doi:10.1080/13607863.2013.875124.

McGraw, K. and Wong, S. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Method*s, 1, 30–46.

Moniz-Cook, E. *et al.* (2008). A European consensus on outcome measures for psychosocial intervention research in

dementia care. *Aging & Mental Health*, 12, 14–29. doi:10.1080/13607860801919850.

Nunnally, J. and Bernstein, I. (1994). Validity. In *Psychometric Theory*, 3rd edn. (pp. 83–113). New York, NY: McGraw-Hill.

O'Connor, D., Ames, D., Gardner, B. and King, M. (2009a). Psychosocial treatments of behavior symptoms in dementia: a systematic review of reports meeting quality standards. *International Psychogeriatrics*, 21, 225–240. doi:10.1017/S1041610208007588.

O'Connor, D., Ames, D., Gardner, B. and King, M. (2009b). Psychosocial treatments of psychological symptoms in dementia: a systematic review of reports meeting quality standards. *International Psychogeriatrics*, 21, 241–251. doi:10.1017/S1041610208008223.

Pallant, J. (2005). *SPSS Survival Manual*. Maidenhead, UK: Open University Press.

Spector, A. and Orrell, M. (2006). Quality of life (QoL) in dementia: a comparison of the perception of people with dementia and care staff in residential homes. *Alzheimer Disease & Associated Disorders*, 20, 160–165.

Thomas, D. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27, 237–246. doi:10.1177/1098214005283748.

Ueda, T., Suzukamo, Y., Sato, M. and Izumia, S. (2013). Effects of music therapy on behavioral and psychological symptoms of dementia: a systematic review and meta-analysis. *Ageing Research Reviews*, 12, 628–641. doi:10.1016/j.arr.2013.02.003.

Vink, A., Birks, J., Bruinsma, M. and Scholten, R. (2013). Music therapy for people with dementia (review). *Cochrane Database of Systematic Reviews*, 9.

York, E. (1994). The development of a quantitative music skills test for patients with Alzheimer's disease. *Journal of Music Therapy*, 32, 280–296.

York, E. (2000). A test-retest reliability of the Residual Music Skills Test. *Psychology of Music*, 28, 174–180. doi:10.1177/0305735600282006.