Lok Lam Daniel. Lai BSc Hons MSc

WORKING WITHIN THE PREFERENCE OF PEOPLE WITH DEMENTIA

Section A: Preferred Music Interventions' Effect on Psychological and Behavioural Outcomes for People with Dementia: A Systematic Review
Word Count 7629 (668)

Section B: Development of the Analysis Scale of Engagement (VASE) Tool for People with Advanced Dementia
Word Count 7998 (81)

Overall Word Count
15627 (749)

A thesis submitted in partial fulfilment of the requirements of
Canterbury Christ Church University for the degree of
Doctor of Clinical Psychology

06 June 2019

SALOMONS
CANTERBURY CHRIST CHURCH UNIVERSITY

Acknowledgments

I would like to thank all of the people who generously volunteered their time and opinions to take part in the research and help develop and try out the tool. I would like to thank the staff and residences from Wigmore Hall, in providing me with this opportunity to conduct this research.

I would like to thank my supervisors Professor Paul Camic and Professor Sebastian Crutch for their support, encouragement, and expertise throughout this process. I would also like to thank Julian West and the Music for Life musicians support the research project and providing many meaningful ideas. Thanks must also go to my course manager Linda Hammond for her support. Thank you, Kenny Chin, for bringing clarity to my queries in relation to statistical analysis.

I would also like to thank my mentor and friend Professor Claudia Lai for her support, help and encouragement throughout the three years of the course. A big thank you to my brothers and sisters from 3G, CCIL, Lingfai and overseas fellowship in your support, love and prayers.

I would like to say thank to my family, Mum, Dad and Lily for their love, patience, prayers and encouragement throughout. Most importantly, I would like to praise the Lord Jesus Christ for His grace, unconditional love and blessings in making all this possible. I thank Him for giving me the opportunity to know everyone who have supported me throughout. May His blessings be upon those who have been part of this memorable journey, Amen.

MRP portfolio summary

Section A is a literature review investigating and exploring the effectiveness of preferred music intervention for people with dementia. It goes on to review the relevant empirical literature on the intervention's effect on people with dementia's behavioural and psychological outcomes. The report examined twelve studies that investigated outcomes including agitation, anxiety, depression and overall emotional state. Inconsistent results are found across studies of most outcomes, indication further research in this area is required. Clinical implication and directions for future research are identified.

Section B is a study that aimed to develop and validate a tablet computer-based observation tool that could continually appraise moment-by-moment changes in the level of engagement of people with advanced dementia during interventions. It discussed the process research team went through to develop the intervention protocol, reporting the testing of its validity and reliability. The results indicate that the current version of the tool has good reliability in some areas. Further investigation and adjustments are needed for the tool to be valid and reliable in measuring the engagement of people with advanced dementia in intervention settings. Clinical implication, the study's limitation and the directions for future research direction are discussed.

Section C is an appendix of supporting materials.

Contents

Section A: Contents

Section B: Contents

Section C: Contents

Section A: Literature Review

PREFERRED MUSIC INTERVENTIONS' EFFECT ON PSYCHOLOGICAL AND BEHAVIOURAL OUTCOMES FOR PEOPLE WITH DEMENTIA: A SYSTEMATIC REVIEW

Word Count 7629 (8297)

A thesis submitted in partial fulfilment of the requirements of
Canterbury Christ Church University for the degree of
Doctor of Clinical Psychology

APRIL 2019

SALOMONS INSTITUTE
CANTERBURY CHRIST CHURCH UNIVERSITY

Abstract

A previous review concluded that preferred music is an effective intervention for managing agitation in people with dementia; however, the majority of the research articles included were pilot studies. A systematic review is carried out in this work to investigate the effect of preferred music on people with dementia. Six search engines yielded 240 papers, of which 12 were eligible. The report reviewed studies that investigated outcomes including agitation, anxiety, depression and overall emotional state. Inconsistent results are found across studies of most outcomes. Methodological issues mean that some studies are prone to bias. Consequently, it is difficult to draw conclusions from the review. The results suggest the need for further investigation into this area of research.

Keywords: dementia, music, singing, agitation, anxiety, depression

Introduction

The World Health Organization has estimated that around 50 million people globally are living with a diagnosis of dementia (WHO, 2018). By 2050 this number is expected to have increased to over 135 million (Alzheimer's Research UK, 2013). In the United Kingdom (UK), data from the National Health Service (NHS) has shown that there are around 850,000 people currently living with dementia, with a yearly cost of care at approximately £26 billion to the UK society (Alzheimer's Society, 2014, 2018). With dementia increasingly affecting more people's lives, it is essential to identify appropriate and cost-effective interventions for this population.

There are different subtypes of dementia; Alzheimer's disease, vascular, Lewy body and mixed-type dementia are the main subtypes. There are also rare forms that are more likely to produce young onset, and these include familial Alzheimer's disease, frontotemporal dementia, familial frontotemporal dementia, posterior cortical atrophy and primary progressive aphasia (Rare Dementia Support, n.d.). As well as the common symptoms of memory loss and cognitive impairment, difficulties can also include psychosocial problems, such as agitation (Palm, Sorg, Armin, Gerritsen & Bernhard, 2018), depression (Kuring, Mathias & Ward, 2018) and anxiety (Kuring, Mathias & Ward, 2018; Orgeta, Qazi, Spector & Orrell, 2015). Other symptoms of rare dementias include changes to personality, behaviour, diet and language (Alzheimer's Research UK, 2018; Rare Dementia Support, 2018). Of these symptoms, agitation is found to be the most common in PWD; it includes different behaviour problems, including repetitive actions, restlessness, wandering, and aggressive behaviours (Pedersen, Andersen, Lugo, Andreassen & Sutterlin 2017).

Research has found that around 80 percent of people with dementia (PWD) report some of the above difficulties, which are often categorised as the behavioural and psychological symptoms of dementia (BPSD) (Margallo-Lana et al., 2001). Dementia is a

progressive condition and cures have yet to be found for any of its subtypes; as a result, it is important to identify interventions that can help manage these life-long difficulties that dementia creates for the individual. Currently, non-drug interventions have focused on promoting the individual's quality of life and reducing the psychological distress that dementia creates for the individual as well as their caregivers.

Psychosocial interventions

Apart from pharmacological treatments that are thought to slow the progression of dementia (NHS, 2017, June 17), it has also been suggested that non-pharmacological interventions (NPI) provide a meaningful and preferred method of care for people with dementia (PWD) (Cabrera et al., 2015). These interventions include cognitive stimulation therapy (Orrell et al., 2017), talking therapies (Cheston & Ivanecka, 2017) and music and art therapies (Raglio et al., 2015; Deshmukh, Holmes & Cardno, 2018). Previous reviews have concluded that NPI provide a useful and cost-effective approach to promote quality of life for PWD (Olazaran et al., 2010).

Music and dementia care. According to the *Dementia Guide: Living Well after Diagnosis*, published jointly by the Alzheimer's Society and the National Health Service (Alzheimer's Society, 2018), music intervention is one of the cost effective, non-drug treatments recommended for dementia care. Music intervention is an umbrella term that described treatments that use "music for a patient during a single episode of care to produce outcomes that were achievable during that session of music" (Evans, 2001, p.9).

During the Second World War, music was found to have positive physical and psychological impacts on veterans (Davis & Hadley, 2015). Since then, health professionals have been exploring the benefits of music for different populations, including PWD (Nilsson, 2008). Reviews on music intervention for dementia patients have determined that it has potential benefits in reducing behavioural disturbance and improving mood (Brotons, Koger

& Pickett-Cooper, 1997; McDermott, Crellin, Ridder & Orrell, 2012; Koger et al., 1999; Nilsson, 2008; Ridder, 2005; Vink et al., 2003, 2011). However, evidence is still somewhat inconclusive about its effectiveness due to the studies in this area generally not being particularly rigorous, or mostly using small sample sizes (McDermott et al., 2012). Notably, however, two early reviews that investigated different types of music intervention concluded that the use of participant preferred music is an important element for managing problem behaviours in PWD (Brotons et al., 1997; Lou, 2001).

It has been proposed that music intervention in dementia care theoretically contributes to a sense of personhood; this is a psychosocial theory that Kitwood (1997) developed, which details the importance of helping to support the identity of a person regardless of their cognitive ability. Personhood is key to the quality of life for a PWD and it is proposed as a key conceptual component in the effectiveness of singing and music groups for this population because group singing activities can support social inclusion and cohesion and, arguably, increase agency whilst providing emotional support (Unadkat, Camic & Vella-Burrows, 2017).

It has been emphasised that maintaining social interconnectedness and relationships is important for PWD (Bidewell & Chang, 2011). This is because PWD have been found to experience a reduction in intimacy and relationship quality, as well as a lack of communication after diagnosis (Unadkat et al., 2017). Music intervention has been found to re-establish communication and build inter-relationships through the use of music between group members, as well as with therapists (McDermott, Orrell & Rider, 2013). Music intervention has also been found to foster a sense of belonging (meaningful relationships), security (feeling safe), fulfilment and purpose (achieving meaningful values), which are thought to be essential factors in Nolan et al.'s (2002) sense framework for promoting good quality of care (Camic, Williams & Meeten, 2013).

Preferred music interventions

Preferred music intervention is defined as a systematic way of presenting music that has been integrated into a person's life and is based on personal music preferences (Gerdner, 1992; Sung & Chang, 2005). The use of preferred music for PWD was first proposed by Gerdner (1997). The intervention was based on previous evidence where listening to one's favourite music (preferred music) can decrease levels of cortisol, which reduces anxiety and promotes relaxation (Chlan, 1998). Gerdner (1997) believed that preferred music influences PWD in two particular ways. Firstly, music acts as a non-verbal communication for PWD, who often experience a decrease in their ability to understand verbal communication. Secondly, music that means something to an individual helps to activate past memories (before the onset of cognitive impairment) that are often found to have remained intact for PWD (Burnside, 1988; Randall, 1991). This has also been observed in situations where music has assisted in creating a sense of familiarity, even in new environments (Song et al., 2002).

Based on these principles, Gerdner (1997) developed a mid-range theory for managing PWD experiencing agitation, which is known as the Individualised Music Intervention for Agitation (IMIA). IMIA proposes that due to cognitive impairment PWD have a lower stress level threshold and a heightened potential for anxiety (Gerdner, 1992). As dementia progresses, the threshold decreases as well, resulting in greater likelihood of heightened levels of anxiety and agitated behaviours (Gerdner, 2012; Buckwalter & Hall, 1987). By using music that is carefully selected based on the individual's preferences before the onset of cognitive impairment, Gerdner (1997) proposed that selected music could act as a way to communicate with the agitated individual. It could also stimulate their remote past memory, shifting the attention away from a possibly confusing and meaningless environment in the present to a more interpretable stimulus (Gerdner, 2012). To select music systematically Gardner (2005) designed the Assessment of Personal Music Preference

(APMPQ), which interviews both PWD and their careers about the PWD's music preference. Overall, studies have confirmed the effectiveness of using preferred music in managing agitation among PWD (e.g. Gerdner, 2000; Regneskog, Asplund, Kihlaren & Norberg, 2001; Sung, Chang & Abbey, 2006a).

A literature review was previously completed examining the use of preferred music to decrease agitated behaviours among PWD (Sung & Marie, 2005). The paper reviewed eight studies between 1993 and 2005. The researchers found consistent results across most studies and concluded that preferred music has a positive effect on reducing agitated behaviour. However, the majority of the research articles that Sung and Marie (2005) reviewed were pilot studies that engaged small samples. Four of the studies recruited a sample size of five or less participants, and only two had a simple size of more than 18. Consequently, they acknowledged that drawing conclusions from these studies should be done with caution and there was a need for further research in this area. In addition, it is also worth noting that although Sung & Marie (2005) made comments on the methodological designs of their included studies, a validated quality appraisal tool was not employed, raising questions on whether the studies were critically appraised. Despite the review's limitations, it was recommended that future reviews should investigate the intervention's effect on other behavioural and psychological outcomes. Since the publication of Sung and Marie's (2005) review, additional studies have been published on preferred music interventions and the use of preferred music for PWD has been extended beyond the use of managing agitative behaviours. Considering the combination of Sung and Marie's (2005) review, as well as their recommendations meant that a more up to date review would be beneficial for the understanding of preferred music for PWD.

Aims of the present review

The aims of this review are to explore the psychological and behavioural variables of anxiety, depression, overall emotional state and agitation in response to preferred music as an intervention for PWD. The review also evaluates the methodology of the studies, discussing the implications for future research and how findings can contribute to clinical practice. Unlike Sung & Marie's review, the present review explores psychological and behavioural factors beyond only reviewing PWD's agitation. Pilot studies are not included because pilot studies often do not have meaningful effective size estimates and they are designed only to inform and evaluate the feasibility of a larger scale study. Therefore, the results might not be generalisable beyond the criteria of the pilot design (Leon, Davis & Kraemer, 2011).

To improve Sung and Marie (2005) review the present research initially intended to compare the effects of non-preferred music interventions with preferred music. However, after a careful investigation of the available literature, it was found that there was a large quantity of publications of non-preferred music interventions. Robb et al. (2018) found 850 publications in their systematic review about music interventions and concluded that meaningful interpretation and cross-study comparisons were difficult. To prevent this issue occurring in the present review it was decided to focus solely on the effects on non-preferred music to enable a more detailed discussion of the intervention. Furthermore, the review took the position of viewing preferred music interventions as a separate intervention to other music interventions, as it has its own theoretical model and rational (Gerdner, 1997), differing from other non-preferred music interventions.

Methods

A systematic search was conducted to explore existing research into the use of preferred music with PWD (Grant & Booth, 2009). An initial search was conducted in relation to preferred music interventions for this population. Online databases, including PsycInfo, ASSIA, Medline, PubMed, Cinahl and Cochrane Library, were searched; only peer-reviewed journal articles were included. The search terms are presented in Table 1. Reference lists were also screened to avoid missing any relevant studies. The search process is outlined in Figure 1.

Table 1. *Table of initial search results*

DATABASE	PSYCINFO	ASSIA	MEDLINE	PUBMED	COCHRANE	CINAHL
Search 1	(dement OR Alzheimer)					
Results	12724	8606	10559	105336	8808	18,825
Search 2	(individualized music OR individualised music OR preferred music OR favoured music)					
Results	181	2151	113	581	618	353
Search 3	(Search 1 AND Search 2)					
Results	4	139	24	19	24	118

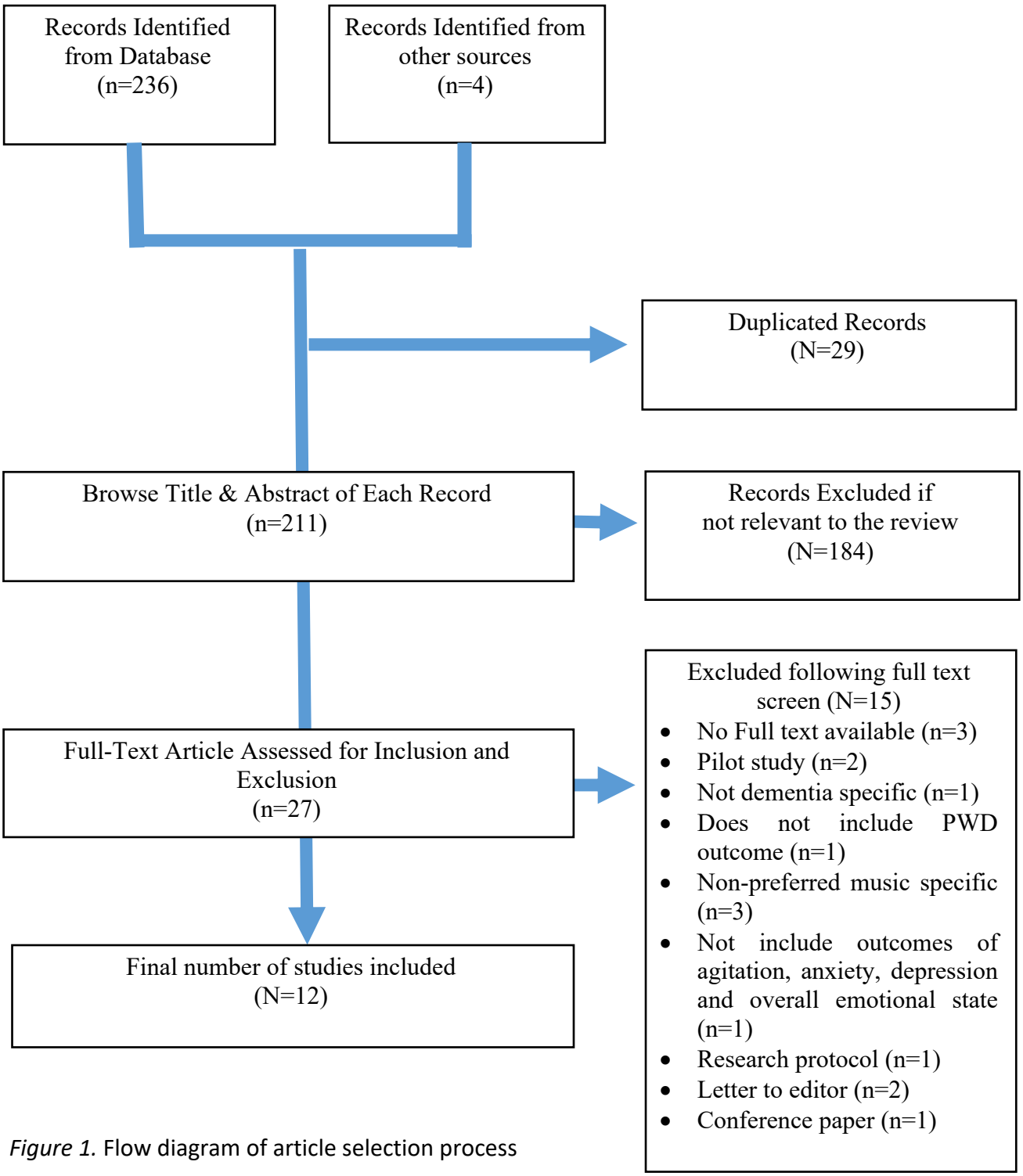


Figure 1. Flow diagram of article selection process

The Mixed Methods Appraisal Tool (MMAT) checklist (Hong et al., 2018) was used as a guide for quality appraisal (Table 3). This is a modified version of the previous MMAT checklist (Pluye et al., 2011). MMAT was selected because it is able to assess different types of research designs, appraising qualitative research, randomised controlled trials (RCT), non-randomised studies, quantitative descriptive studies and mixed methods studies. Unlike the earlier version, the current version appraises methodology in five categories, but it no longer encourages users to calculate an overall score (Hong et al., 2018). Table 2 and 3 presents the information extracted from each study and MMAT appraisal.

Inclusion and exclusion criteria

Inclusion criteria.

1. Adapted from Sung and Marie (2005), the criteria specific to the topic were: (a) intervention studies of preferred or individualised music; (b) people with Alzheimer's disease or other dementias.
2. Only peer-reviewed journals.
3. Measures of the psychological or behavioural outcome of anxiety, depression, overall emotional state and agitation for PWD.

Exclusion criteria.

1. Studies that used music interventions, but did not involve participating in preferred music intervention.
2. Pilot studies.
3. Studies that did not discuss strategies used to identify PWD's preferred music; studies that only stated that preferred music was used, without further explanation, were not included as this would offer little insight whether music was "systemically selected" (Gerdner, 2012), which is a criteria for the present review.

4. Studies without an intervention component (e.g. protocol development, assessment and publications that only discuss theory).
5. Studies without an explicit methodology.
6. Studies published in a language other than English.

Results

The 12 studies included in the research were carried out internationally – including North America, Europe and Asia – by researchers from different professional backgrounds. This included nurses, social workers, music therapists and psychologists.

Study samples

The participants in the included studies were older people (age range between 65 and 100) diagnosed with different forms of dementia. The majority of the studies took place in long-term care facilities, with one completed in a community centre (Sanchez et al., 2016). All studies used screening measures to assess cognitive impairment, including the Global Deuteration Scale (GDS) (Cheung, Lai, Wong & Leung, 2018; Gerdner, 2000; Sanchez et al., 2016; Sung, Chang & Lee, 2010; Ledger & Baker, 2007), Mini-mental State Examination (MMSE) (Cooke, Moyle, Shum, Harrison & Murfield, 2010; Garland, Eppingstall & O'Connor, 2007; Guetin et al., 2009; Ledger & Baker, 2007; Raglio et al., 2015), Clinical Dementia Rating (CDR) (Raglio et al., 2015; Sakamoto, Ando & Tsutou, 2013) and the Cognitive Function Scale (CFS) (Thomas, Baier, Josar, Ogarek, Trepman & Mor, 2017). Three of the studies recruited people with moderate levels of impairment (Cooke et al., 2010; Guetin et al., 2009; Ledger & Baker, 2007), two recruited those with severe levels (Garland et al., 2007; Sakamoto et al., 2013), and five studies included both moderate and severe levels (Cheung et al., 2018; Gerdner, 2000; Sanchez et al., 2016; Sherrat et al., 2004; Sung et al., 2010). One study included those with mild to severe levels (Raglio et al., 2015) and two studies did not report severity levels (Kwak et al., 2018; Thomas et al., 2017). The studies

engaged a broad sample size, ranging from 22 to 6,298 participants, with the majority of the participants being female.

Thomas et al's (2017) is the only study that adopted a sample size larger than 130 participants. This study recruited 98 facilities in the US that practiced the preferred music intervention of Music & Memory (M&M). Outcome measures of agitation, depression and medication use were compared with "matching comparison" sample. The matching was done by a Medicare software program that matches participants to PWDs of other facilities that are in the same geographical regional location, with of similar age and functioning (e.g. cognition and activity of daily living). It is worth noting that all outcomes were collected through Minimum Data Set (MSD), a national database, and as a result researchers had little or no control on whether the intervention was administered as intended or not. Furthermore, the paper did not mention what researchers did to ensure that outcome measures were recorded appropriately across multiple care homes. As a result, despite the large sample size offering better generalisability, the unclear picture of how data were collected and monitored meant that conclusions were difficult to be drawn from the study. Further discussion about the study's intervention and outcome will be discussed below.

Interventions

Overall, the review found a wide range of preferred music interventions. In this section, the interventions will be described and categorised into two groups: passive music and active music interventions. Passive music interventions are where participants listen to recorded or live music, whilst active music interventions invite participants to actively engage in activities (e.g. singing, playing a musical instrument, and improvising music with a facilitator) (AMTA, 2015).

Passive music interventions. Eight studies used passive music as one of their testing conditions. Of these studies, five used a published protocol, whilst three studies did not use or did not specify a protocol. The earliest developed protocol was the Individualised Music Protocol (Gerdner, 2000; Sanchez et al., 2016), first designed by Gerdner (1996) based on a mid-range theory she had previously developed and validated (Gerdner, 1997). The protocol was initially intended to reduce agitation in dementia patients and it includes a summary of the risk factors for agitation, assessment criteria and assessment of music preference (Gerdner, 2012).

The Music & Memory (M&M) protocol is an adaptation of Gerdner's (1996) Individualised Music Intervention (Kwak, Anderson & Valuch, 2018; Thomas et al., 2017). This intervention was designed to use relatively inexpensive technology, such as an iPod, to create a PWD preferred playlist to help reduce BPSD symptoms. In order to use M&M in clinical settings, at least one member of staff from the facility will have to undertake three 90-minute training sessions. The training includes education about the benefits of using personalised music and the legal aspects of music sharing, as well as how to create a personalised music playlist on iTunes® and how the program should be carried out. Once training has been completed, the certified facilities will also be able to access online resources for support and ideas in relation to implementation. The general guidelines on how to implement M&M are published and can be requested online via: https://musicandmemory.org/training-publications/request-guide/.

Listen to Music (LtM) (Raglio et al., 2015) is a similar intervention to M&M that is also designed to reduce BPSD symptoms among PWD. LtM is a 30-minute music listening program where PWD are asked to listen to music from their preferred playlist in a quiet room, without headphones or any interaction with a music therapist or formal caregiver (Raglio et al., 2015).

Guetin et al. (2009) used a music therapy technique that is known as the U technique. This is a technique created by the record publishing company Music Care. Based on PWD's choice of music, a computer programme creates a 20-minute sequence for the PWD to listen to. The sequence often begins with songs with stimulating rhythms. Then, by gently reducing the rhythms, orchestral formation, frequency and volume, the sequence will bring the PWD into a more relaxed state; towards the end of the sequence, the songs will then become more stimulating again.

A variety of equipment was used by PWD in the researches to listen to music, including CD players (Gerdner, 2000; Sakamoto, Ando & Tsuto, 2013; Sanchez et al., 2016; Sung, Chang & Lee, 2010), iPods (Thomas et al., 2017; Kwak, Anderson & Valuch, 2018) and audio speakers (Raglio et al., 2015). The majority of the studies did not offer any clear rationale as to why they used specific equipment to deliver the music, except for Thomas et al. (2017), who spoke about the iPod as a cost-effective technology.

Active music interventions. Ledger & Baker (2007) and Raglio et al. (2015) both employed music therapy as one of their interventions. The British Association of Music Therapy defines music therapy as "an established psychological clinical intervention, which is delivered by HCPC registered music therapists to help people whose lives have been affected by injury, illness or disability through supporting their psychological, emotional, cognitive, physical, communicative and social needs" (BAMT, 2017). This type of therapy might involve a music therapist using different techniques, including singing, changes in melody and rhythm, as well as using different musical instruments to establish a relationship with the PWD (Raglio et al., 2015). The aim is to facilitate the expression of emotions and promote moments of "attunements".

The Music-with-movement intervention (Cheung, Lai, Wong & Leung, 2018) is an intervention recommended for people with moderate-stage dementia; it encourages people to

move their larger body muscles while their preferred music is played on a music player. The intervention does not require any fine motor muscle movements or verbal ability. Cooke et al. (2010) conducted a musician-facilitated live group music program for PWD, which included 30 minutes of musician-led singing using a guitar and 10 minutes of pre-recorded instrumental music for active listening (Cooke et al., 2010).

Selecting preferred music. When selecting preferred music for PWD, four of the studies used validated questionnaires to gather PWD's choice of songs to ensure they would have personal meaning. One of the questionnaires was the Assessment of Personal Music Preference (APMPQ) (Gerdner, 2000; Sanchez et al., 2016). The APMPQ was initially developed and tested by Gardener (2005) and it has two versions: one designed for PWD, and the other for caregivers. The caregiver version is only used as a proxy measure when PWD do not have the capacity to make self-selections. Both versions ask nine similar questions that aim to obtain information about preferred types and forms of music, favourite singers and music that impacted their lives before dementia onset.

The Music Preference Survey (MPS) was used by two of the reviewed studies (Cooke et al., 2010; Sung et al., 2001) and it is a modified version of the APMP (Gerdner, 2000). This is a brief tick-box questionnaire that asks respondents to choose from the categories provided under the favourite types of music, forms of music and favourite artists. The MPS can also be filled in by the caregiver if the PWD is no longer able to report this information.

Some studies chose to select preferred music lists after interviewing both PWD and caregivers (Garland et al., 2007; Raglio et al., 2015; Sakamoto et al., 2013; Thomas et al., 2017; Kwak et al., 2018), whilst others interviewed only the PWD (Gutin et al., 2010; Ledger & Baker, 2007) and one only interviewed caregivers (Cheung et al., 2018). When choosing music, as well as asking PWD's preferences, all studies stressed the importance of identifying songs that were significant to that person's life experience

(Sanchez et al., 2016). Some identified the importance of reviewing the song preference throughout the intervention (Cheung et al., 2018; Cooke et al., 2010). Choosing and changing songs might also involve close observation of participants' behaviour and making a clinical judgement to determine the period of their life that was remembered or spoken about most frequently (Sakamoto et al., 2013).

Study design

Table 2 provides information about the data extracted from each study. The MMAT (Hong et al., 2018) critical appraisal tool was used to appraise the studies' methodological quality (see Table 3). The reviewed studies adopted various research designs. Eight of the studies employed randomised controlled trials (RCT) (Cheung et al., 2018; Cooke et al., 2010; Sanchez et al., 2016; Guetin et al., 2009; Kwak, Anderson & Valuch, 2018; Raglio et al., 2015; Sakamoto, Ando & Tsutou, 2013), two involved repeated measure designs (Ledger et al., 2007; Sung, Chang & Lee, 2010; Thomas et al., 2017), and one employed a crossover repeated measure design where PWD served as their own control (Gerdner, 2000). With any uncontrolled design there is a greater potential for bias and error, such as practice effect among participants (Thiese, 2014). Therefore, evidence from these uncontrolled trials, although valuable, should be interpreted with caution.

The RCT is considered the "gold standard" in research design (Akobeng, 2005). However, based on the MMAT appraisal there was found to be different levels of quality across the RCTs. For example, four of the studies did not explicitly discuss how they monitored the participation of PWD in the intervention (Cheung et al., 2018; Cooke et al., 2010; Raglio et al., 2015; Sakamoto et al., 2013). This means that these studies might potentially include participants who did not receive the full course of treatment.

Blinding was not undertaken in any of the studies. The fact that the participants were aware of their allocated intervention group might have affected their response to the outcome assessment and increased the risk of performance or social desirability bias. For three of the studies, the blinding of the assessor was either not carried out (Kwak et al., 2018; Sung et al., 2010), or it was unclear whether the assessor that conducted the data collection was blinded or not (Sanchez et al., 2016). Only Sung et al. (2010) explained that blinding was not possible due to the nature of the study design. Inability or failure to blind the outcome assessor meant that these studies might be prone to detection bias.

Sakamoto et al. (2013) randomised care homes as clusters instead of individual participants. This meant that the studies were more prone to sampling bias. The significant difference found in age and anxiety between the two groups at baseline might reflect such bias and suggests that the baseline data might not be comparable. In the case of Ledger et al. (2007), they also acknowledged that confounders, such as participants' agitation levels over the year, illnesses, hospitalisations, changes in medications, bedroom changes, and deaths among family and friends, were not considered or accounted for.

Table 2. *Data extracted from reviewed studies*

Papers	Participants	Location	Intervention	Design	Data collection	Measurement for preferred music	Measures	Psychological impact finding(s)
Cheung et al., (2018)	165 PWD - 40 male, 125 female Deteriorati on Scale stage 5-6	Hong Kong, 12 care homes	Duration:1 hr weekly for 12 weeks - Music listening (n = 54) - Music-with-movement intervention (n= 58) - Social Activities group (n=53)	Multi-centre RCT	- Baseline, - mid-point (6 weeks) - immediate post-intervention	Interview	Anxiety - RAID Depression - GDS	Anxiety – - Insignificant group by time effect - Significant reduction of anxiety in both the MM and ML groups between baseline and at post-intervention Depression - - MM group showed a significant reduction in depressive symptoms between baseline and post intervention but not and not the ML
Cooke, et al., 2010	47 PWD - 14 male, 33 female. Age range 75-94. MMSE mean 16.51	Queensland, Australia, 10 care homes	Duration: 40 mins, three times per week for 6 months - Active Music group (live singing and Prerecording listening) (n=23) - Reading control group. (n=24)	Randomized cross-over design. Changed over after 5 weeks of wash out period.	- Baseline, - mid-point (4 weeks) - immediate post-intervention	MPS, review after the first three sessions	Agitation – CMAI Anxiety - RAID	Agitation- - No significant overall effect of the music programme in ameliorating agitation - Significant increase in the frequency of verbal aggression over time, regardless of group ($F(2,46)=3.534$, $p<.05$). Anxiety- - No significant overall effect of the music programme in reducing anxiety

Garland, et al., 2007	30 PWD - 10 male, 20 female, Age 66-93 (mean - 79). CMAI > 1, MMSE 0- 12 (mean 2.5)	Australia Nine care homes	Duration:15 mins a day - Preferred music listening - Placebo: horticultural text listening - Family stimulated presence - Usual care	Randomised Crossover design.	- Before intervention - During intervention - 15 minutes after intervention	Questionnaire Interview	Agitation - CMAI	Physical Agitation – - No significant different between stimulate presence and music found during the intervention. - Music performed better than usual care ($F(1,29)=4.67$, $p=.039$), but not placebo. - Stimulated presence performed better than placebo ($F(1,29)=8.29$, $p=.007$) and usual care ($F(1,29)= 10.2$, $p=.003$), but not music. Verbal agitation - No significant different between stimulate presence and music found during the intervention. - Significant difference between stimulation presence and placebo were found ($F(1,29)=4.78$, $p=.037$). - No significant between music with placebo and usual care conditions were found with music

Gerdner, 2000	39 PWD – Male, 9 female, 30 Mean age 82 (range 70-99), GDS scores 3-7, mean 6).	Australia, 6 long-term-care facilities	Duration: 30 mins, twice weekly for 6 weeks - Individualized music listening (n=16) - Classical relaxation music (n=23)	Cross over repeated measures. Change over after 2 weeks "washout" period)	- Baseline - Immediate post-intervention	APMPQ	Agitation - CMAI	Agitation – - Significant two-way interaction between phases ($F(2, 74)=32.93$, $p=.001$) and minutes $F(5,2763)=53.28$, $p=.001$) - Significant reduction in agitation during and after preferred music compared to classical relaxing music ($p <.0001$). - Significant reduction in agitation during and after preferred music compared to classical relaxing music ($p <.0001$).
Guetin et al, 2009	30 PWD - 8 male, 22 female. age (range 70-95), MMSE score between 12-25.	France one care home	Duration: 20 mins, once weekly - U method music listening (n=15) - Reading group (n=15)	RCT.	- Baseline, - Week 4, - Week, 8, - immediate post-intervention (Week 16) - Follow up (Week 24).	Questionnaire Interview	Anxiety - Hamilton scale Depression - GDS	Anxiety – - Significance difference in all time point between group ($p < 0.001$) - Decrease means score for music group overtime ($p < 0.001$). - Persistence effect of music therapy were found in W 24 ($p=.002$). Depression – - Significant difference of the 2 group were found at all time point ($p < 0.05$) - Decrease in depression for music group.. Persistent effect on depression was found by both groups.

Kwak, et al., 2018	59 PWD - 13 male, 46 female. with dementia diagnosis	U.S.A., 10 care homes	Duration: 14 weeks, listen when needed (including 2 weeks washout period before crossing over) - M&M personalize playlist on ipod) - treatment as usual (no music)	RCT crossover design.	- Baseline, - Week 6, - Week 8 - immediate post-intervention (Week 14)	M&M protocol; Interview caregivers and PWD	Agitation - CMAI Medication – standardized form NPI – irritability subscale Depression – NPI depression subscale Overall emotional state- NPI scale - Irritability - Disinhibition - Depression	Agitation – - No significant difference were found between conditions overtime Medication- - No significant effect observed Depression- - No significant difference were found between conditions over time - Inconsistent result between conditions of the two phases - Depression- significant condition by time interaction for NPI depression found, suggesting decreased in depression on M&M but increase in depression in treatment as usual overtime ((F(1, 9)=5.42, p=.04) Overall emotional state – - NPI global scale was not reported

Ledger & Baker, 2007	45 PWD - 5 male, 40 female. Recruited from GDS 4-5, MMSE < 23 or MSQ >2). Age range (71-100)	Queensland and Victoria, Australia;13 care homes.	Duration: 30-45 minutes, Weekly for 1 year (no less than 42 weeks) - Music therapy group (n=26) - Treatment as usual (n=19).	Longitudinal repeated measure design	- Baseline - three-months, - six months, - nine months - immediate post-intervention (12 months)	Participants choosing or requesting favourite songs in the group	Agitation – CMAI	Agitation – - No significant different agitated behaviour over time. - Total CMAI means for both groups fluctuated from one data collection point to the next.
Raglio et al, 2015	120 PWD - 26 male, 94 female. Age 65>; CDR range 1-4; MMSE <18; MPI <18.	Italy; 9 care homes.	Duration: 30 minutes, Twice weekly for 10 weeks - Listen to Music (LtM) - playlist on speaker Active (n=40) - Music therapy - nom preferred music specific (n=40) - Standard of care (SC) group (n=40)	RCT	- Baseline - immediate post-intervention - 2 months after intervention	Music therapist created playlist after interview	Depression - CSDD Overall emotion state - NPI	Depression - - No significant between group observed - All groups showed reduction overtime in CSDD (P = .001) Overall emotion state - No significant between group observed NPI global score - All groups showed reduction overtime in (P ≤ .001) - Trend where the NPI global score fell 28% in the MT group, 12% in the LtM group, and 21% in the SC group at the end of treatment.

Sakamoto, et al., 2013	39 PWD. Age 65>; CDR = 3. Gender (not reported)	Kobe City, 4 care homes and a special dementia hospital	Duration: 30 minutes weekly between 10am-11am, for 10 weeks - Interactive music intervention by music facilitator (n = 13) - Music - listening (n = 13) - listed to selected music via CD player - non-music carer interaction control (n = 13)	RCT	Short term effect - 5 minutes prior intervention - 5 minutes after intervention Long Term effect (10 weeks) - Baseline - immediate post-intervention - Three weeks after intervention	Interviewing Analysed participants' behaviour to determine the period of their life they remembered	Short term: Emotional state - The faces scale; Heart rate measurement before and after. Long term: (10 weeks) Agitation - BEHAVE-AD	Short term - Emotion state The faces scale - Interactive group exhibited an even greater improvement in the emotional state among the three groups (Z –3.2, $p < 0.01$). - Patients in the Passive group were in a significantly more comfortable mood after the intervention compared with before the intervention (Z –2.3, $p < 0.01$) Heart Rate - Passive and interactive music interventions caused short-term parasympathetic dominance (reduced stress), $F(1, 36) = 4.968$, $p < 0.01$; HF, Z –2.6, $p < 0.01$), but not in control group. Long term - Agitation - Significant interactive of interactive intervention, compared with passive music intervention and a no-music control condition - Affective disturbance, (Z –2.3, $p < 0.025$) - anxieties and phobias, (Z –2.3, $p < 0.025$) - Paranoid and delusional ideation, (Z –2.7, $p < 0.025$) - Aggressiveness (Z –2.6, $p < 0.025$) Activity disturbance, (Z –2.5, $p < 0.025$).

Sanchez et al, 2016	22 PWD - 15 male, 7 female. Mean age 88.41 (range77-102), GDS scores 6-7	Coruna, Spain. Specialized dementia elderly centre.	Duration for: two 30 minutes weekly for 16 weeks - Individualized music intervention (n=11) - MSSE (n=11)	RCT	- Baseline, - mid-point (8 weeks) - immediate post-intervention - 8 weeks follow up	APMPQ	Agitation - CMAI Depression- CSD Anxiety- RAID	Agitation – - No significant changes were found in both group at pre-, mid-, post-. And follow up. But trend of improvement in score at follow-up in both groups, were found. Depression – - No significant change for both group overtime - Anxiety – - Improvement in mood at follow-up for MSSE condition, but not individualized music $F(1,16)=9.822$, $p=0.006$).
Sung, Chang & Lee, 2010	52 PWD. Age 65>; Global Deteriorati on Scale 4-6;	Taiwan. Two care homes.	Duration: 30 mins, twice weekly in mid-afternoon for 6 weeks - Preferred music listening vis CD player (n=29) - usual care (n=23)	Quasi-experimental Repeated measures	- Baseline - Immediate post-intervention	MPS	Anxiety - RAID	Anxiety - Significantly lower anxiety score at six weeks compared with those who received the usual standard care with no music ($F = 12.15$, $p = .001$)

Thomas, Baier, Josar, Ogarek, Trepman & Mor, 2017	6298 PWD at M&M facilities and 6278 PWD matching comparison facilities. Age 65>, ADL <24, CFS > 2	U.S. 98 Care facilities trained in M&M. Comparison facility were match in exact geographical region. Matching were done using Medicare 5-Star compare rating	Durations: 1 year - M&M (personalize playlist on ipod) (n=6298) - Usual care (n=6278)	Repeated measure	- pre-intervention-care as usual (2012) - Post-intervention-Post intervention (2013)	M&M protocol Interview caregivers and PWD	Antipsychotic and Anxiolytic use – Minimum Data Set (MSD), Agitation - Aggressive ABS. Depression - PHQ-9	Antipsychotic – - Significant different in PWD discontinuing antipsychotic found between the two condition facilities overtime (p =.04) decrease of antipsychotic used increased in M&M facilities from 17.6% to 20.1% post-M&M intervention, oppose to 15.9 to 15.2 among comparison facilities. Anxiolytic medication – - Significant difference in PWD discontinuing anxiolytic found between two group (p=.03) post-intervention, decrease of anxiolytic used increased in M&M facilities from (23.5% to 24.4%) oppose to comparison facilities (24.8% to 20 %) . Agitation – - Significant different of staff reported agitation between M&M facility and comparison facilities (p=.04), from 50.9 to 56.5 in the M&M facility to 55.8%-55.9% in the comparison facilities. Depression – - No significant different were found between M&M and comparison facilities.

*ABS = Abe's Behavioral and Psychological Symptoms of Dementia score, MMSE = Mini Mental State Test, GDS = Geriatric Depression Scale.. CMAI = Cohen-Mansfield Agitation Inventory, M&M= Music & Memory, ADL = Activities of Daily Living, PHQ-9= Patients Health Questionnaire-9, APMPQ = Assessment of Personal Music Preference Questionnaire, NPI= Neuropsychiatric Inventory, RAID = Rating Anxiety in Dementia, APMPQ = The Assessment of Personal Music Preference Questionnaire, CSDD = The Cornell Scale for Depression in Dementia, BEHAVE-AD = Behavioural Pathology in Alzheimer's Disease, RCT = Randomised Control Trial, MSSE = Multisensory Stimulation Environment

Table 3. *Mixed Methods Appraisal Tool (MMAT) checklist*

RANDOMIZED CONTROLLED TRIALS								Comment
	S1. Are there clear research questions?	S2. Do the collected data allow to address the research questions?	2.1. Is randomization appropriately performed?	2.2. Are the groups comparable at baseline?	2.3. Are there complete outcome data?	2.4. Are outcome assessors blinded to the intervention provided?	2.5 Did the participants adhere to the assigned intervention?	
Cheung, Lai, Wong & Leung, 2018	Yes	Yes	Yes	Yes	Yes	Yes	Cannot determine	Cannot tell whether participant did adhere to the treatment or not
Cooke, Moyle, Shum, Harrison & Murfield, 2010	Yes	Yes	Yes	Yes	Yes	Yes	No	Some participants did adhere to the treatment
Garland, Eppingstall & O'Connor, 2007	Yes	Yes	Yes	Yes	Yes	No	Yes	Assessors are not blinded for data collection
Guetin et al, 2009	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A
Kwak, Anderson & Valuch, 2018	Yes	Yes	Yes	Yes	No	No	Yes	Assessors are not blinded for data collection NPI global scale score was not reported
Raglio et al, 2015	Yes	Yes	Yes	Yes	Yes	Yes	Cannot determine	Cannot tell whether participant did adhere to the treatment or not
Sakamoto, Ando & Tsutou, 2013	Yes	Yes	Yes	Cannot determine	Yes	Yes	Cannot determine	Study did not report baseline statistic Cannot tell whether participant did adhere to the treatment or not
Sanchez et al, 2016	Yes	Yes	Yes	Yes	Yes	Can't tell	Yes	Cannot tell if assessor was blinded in the study or not

NON-RANDOMIZED STUDIES								Comment
	S1. Are there clear research questions?	S2. Do the collected data allow to address the research questions?	3.1.Are the participants representative of the target population?	3.2. Are measurements appropriate regarding both the outcome and intervention (or exposure)?	3.3. Are there complete outcome data?	3.4. Are the confounders accounted for in the design and analysis?	3.5. During the study period, is the intervention administered (or exposure occurred) as intended?	
Gerdner, 2000	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A
Ledger & Baker, 2007	Yes	Yes	Yes	No	Yes	No	Yes	Study report in the discussion that some confounding variables not accounted for in the analysis (e.g. changes in medications, bedroom changes, interventions received and deaths among friends and family)
Sung, Chang & Lee, 2010	Yes	Yes	Yes	Yes	Yes	Cannot determine	Yes	Study did not report whether confounding variables were considered in design and analysis or not.
Thomas, Baier, Josar, Ogarek, Trepman & Mor, 2017	Yes	Yes	Yes	Cannot determine	Yes	Yes	Cannot determine	Cannot tell if the Minimum Data Set (MSD) data study used is appropriate as outcome measure or not. As study was drawn from national database, study was not able to control whether, intervention was administered as intended or not.

Outcome measures

Comparing with other non-music interventions. The majority of the studies evaluated the impact of preferred music in comparison to no music. This included treatment such as Multisensory Stimulation Environment (MSSE) (Sanchez et al., 2016), reading (Cooke et al., 2010; Guetin et al., 2009) and social activities (Cheung et al., 2018). Overall, there were no clear and significant findings that indicated preferred music was superior to MMSE or reading in any outcome measured, but there was evidence that it performed better than social activities in improving mood. Some studies also compared interventions with the standard care in the care home (Kwak et al., 2018; Ledger et al., 2007; Sakamoto et al., 2013; Sung et al., 2010) and the findings from the comparisons were inconsistent. These findings will be discussed in greater detail below.

Studies also compared different types of music interventions, including comparing classical relaxing music (non-preferred music) with preferred music listening (Gerdner, 2000), music therapy with music listening (Raglio et al., 2015), and active music intervention with passive music intervention (Cheung et al., 2018; Sakamoto et al., 2013).

Psychological outcomes.

Agitation. The majority of the studies used the Cohen-Mansfield Agitation Inventory (CMAI) to measure changes over time (Cohen-Mansfield, 1986). This inventory consists of 29 items that assess the level of agitation in older people. This is an observation tool where the assessor is asked to rate the frequency of physically aggressive, physically non-aggressive and verbally agitated behaviours (Finkel, Lyons & Anderson, 1992). The scale reports good reliability and validity (Zare, Birashk & Ebrahimi, 2012). One of the studies (Gerdner, 2000) found a significant two-way interaction between different phases (baseline, individualised music, classical relaxation music) ($F(2, 74) = 32.93$, $p = .0001$) and time (per 10-minute

increments) (F (5, 2,763) = 53.28, p = .0001). This work also found that individualised music resulted in a more significant reduction of agitated behaviour compared to classical music (relative to baseline) ($p < .001$). A post hoc comparison test found that the reduction in agitation not only occurred during the 30 minutes intervention period, but it was consistent 30 minutes after preferred music listening at follow-up ($p < .0001$). However, other studies that also used CMAI found no significant reduction in PWD's level of agitation in response to preferred music listening over time (Kwak et al., 2018; Ledger et al., 2007; Sanchez et al., 2016). Ledger et al. (2007) found that the CAMI mean score fluctuated from one data collection point to the next. Kwak et al. (2018) found no significant difference in terms of music condition and care as usual. Cooke et al. (2010) used a modified version – the Cohen-Mansfield Agitation Inventory: Short Form (CMAI-SF) – and they also failed to find an overall effect of the music programme in reducing agitation. Interestingly, a significant increase in the frequency of verbal aggression over time was observed in both the control and music condition (F (2, 46) = 3.534, $p < .05$).

Thomas et al. (2017) used the Aggressive Behaviour Scale (ABS) to measure behavioural disturbance. The ABS rated PWD in terms of their physical, verbal and other behaviours directed toward others, as well as PWD's rejection of care. Each of these domains are scored on a four-point scale (Perlman & Hirdes, 2008). Reduced rates of problematic symptoms and behaviours were found in Music and Memory (M&M) facilities (50.9% behaviour improvement pre-intervention to 56.5% exhibiting improvement post-intervention) when compared with the comparison facilities (55.8% pre-intervention and 55.9% post-intervention) ($z = 2.01$, $p = .04$).

Sakamoto et al. (2013) observed the effect of preferred music listening compared to no music over time using the Behavioral Pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD). They found reductions in affective disturbance (Z –2.3, $p < .025$), anxieties

and phobias (Z –2.3, $p < .025$), paranoid and delusional ideation (Z –2.7, $p < .025$), aggressiveness (Z –2.6, $p < .025$) and activity disturbance (Z –2.5, $p < .025$) in the music group, but not the non-music group. They concluded that music interventions could reduce stress in individuals with severe dementia.

Anxiety. The majority of the studies used the Rating Anxiety for Individuals with Dementia (RAID) as their evaluation tool to investigate the impact of preferred music on anxiety. RAID gathers self-reported information from PWD and caregivers, as well as gathering other relevant clinical information to rate the PWD's level of anxiety (Shankar, Walker, Frost, Orrell, 1999). This is a widely used tool that assesses signs of anxiety over a period of two weeks (Shankar et al., 1999). Assessors are asked to provide a 0–10 score on each of the 18 items of the scale. An overall score of 11 or more indicates possible significant clinical anxiety. The majority of studies that used RAID found no significant indication that preferred music reduced PWD anxiety over time (Cheung et al., 2018; Cooke et al., 2010; Sanchez et al., 2016). Interestingly, Sanchez et al. (2016) found significant reductions of anxiety in those with the multisensory stimulation (MSSE) condition ($F (2, 16) = 2.141$, $p = .013$), but not the individualised music group.

Despite finding an insignificant interaction group-by-time effect, Cheung et al. (2018) carried out a pairwise comparison of anxiety between baseline and at post-intervention. They found a significant reduction of anxiety in both the Music-with-movement ($p < .001$) and the music listening groups ($p = .006$), but not in the social activity group. Interestingly, Sung et al. (2010) found significant reductions of mean RAID scores in the preferred music condition between baseline (mean = 10.93; SD = 5.46) and six weeks into the intervention (mean = 8.93; SD = 4.86) ($t = 5.64$, $p < .001$). ANCOVA also showed that PWD who received preferred music for six weeks had a significantly lower anxiety score than those in the usual standard care condition ($F = 12.15$, $p = .001$).

Guetin et al. (2009) used a self-report measure: the Hamilton Scale of Anxiety. Although not reporting the scale's validity or reliability, results showed a decrease in mean anxiety score over time for the music therapy group ($p < .001$). Significant differences between baseline and follow-up were also reported, suggesting a persistence effect of music therapy on anxiety was observed eight weeks after the intervention ($p = .002$).

Depression. Five studies investigated the effect of preferred music on depression. Two used GDS as one of their outcome measures. Guetin et al. (2009) found a significant difference in GDS score between the music and the control group (reading) at all time points ($p < .05$). They also observed significant improvements in depression ($p < .01$) in the music therapy group from weeks 4–16. Interestingly, Cheung et al. (2018) found significant reductions in depressive symptoms between baseline and immediate post-intervention of the Music-with-movement (active) condition, but not in preferred music listening condition (passive).

Raglio et al. (2015) used a self-report measure – the Cornell Scale for Depression (CSDD) – to assess major depression in PWD; this has good test-retest reliability and high internal consistency (Bradt, Burns & Creswell, 2013). The results demonstrated a significant reduction across all conditions ($P = .001$); this included the preferred music listening (LtM), active music therapy (non-preferred music specific) and control (standard of care) conditions, suggesting that preferred music might not be superior to other care interventions. Like Raglio et al. (2015), Sanchez et al. (2016) used the CSDD, but they found no significant change in depression scores between baseline and immediate post-intervention (16 weeks) of the individualised music group. In fact, an increasing trend from baseline to post-intervention was observed, indicating that participants' "depression" level increased after individualised music intervention. However, a significant reduction in CSDD was found at eight weeks follow-up, and a similar effect was also found in the cognitive stimulation group ($F (1, 16) =$

9.822, p = .006, η2 = .374). It was concluded that the significant improvement observed at follow-up might be an intervention effect, but it could also be due to an accessory effect, such as the seasonal differences between the intervention (winter/spring) and follow-up periods (summer).

Kwak et al. (2018) and Thomas et al. (2017) both evaluated the impact of a preferred music program (M&M) on depression. Kwak et al. (2009) used the NPI depression subscale, whilst Thomas et al. (2017) used the PHQ-9; both are well-validated tools. Both studies found no significant difference between the control condition and M&M over time. Although Kwak et al. (2018) did observe a reduced trend in NPI scores for M&M, it is worth noting that the NPI depression subscale uses only one question, which raises questions about whether the tool is sensitive enough to detect change.

Overall emotional states. Sakamoto et al. (2013) used the faces scale and heart rate activities to monitor emotional state. The faces scale is an observation tool designed to assess people with severe dementia who cannot verbally express their emotions; this involves trained staff evaluating PWD's positive or negative facial expressions. The study did not report reliability or rating procedures. They found that PWD in the passive music listening group were in a significantly better mood post-intervention (Z –2.3, p < .01) than pre-intervention, and PWD in the interactive group exhibited an even greater improvement in their emotional state than the listening and control groups (no-music group); (Z –3.2, p < .01). The study also found that both passive and interactive music interventions caused short-term parasympathetic dominance in heart rate (F (1, 36) = 4.968, p < .01; HF, Z –2.6, p < .01), but this was not found in the control group. Based on these results, researchers interpreted that music intervention promotes a more positive emotional state than a no-music control.

Raglio et al. (2015) was the only study to report the NPI global score, finding a reduction in all groups ($P \leq .001$) at the end of treatment. The global score fell by 28 percent in the music therapy condition (non-preferred music), 12 percent in the preferred music listening condition, and 21 percent in the social care condition. They concluded that there was no evidence to suggest that music interventions might be better at improving mood than standard social care. Despite using the NPI scale, Kwak et al. (2018) did not report the NPI global score.

***Medication*.** Only two studies assessed medication changes as an outcome (Kwak et al., 2018; Thomas et al., 2018). Kwak et al. (2018) observed no significant effect on medication over time and between conditions (M&M versus treatment as usual). Thomas et al. (2018), however, found a significant difference in the number of PWD discontinuing antipsychotics between the facility that used M&M and the facility using the standard care condition ($p = .04$); the M&M facility decreased the use of antipsychotics (from 17.6% to 20.1% post-intervention, as opposed to 15.9% to 15.2% for standard care). Similar findings were found regarding to the use of anxiolytic medication over time ($p = .03$), where a higher percentage of participants stopped using anxiolytics in the M&M intervention facility, from 23.5 to 24.4%, as opposed to 24.8 to 20 percent for comparison facilities. The results suggested that preferred music listening potentially decreases the need for medication in managing PWD's behavioural and psychological symptoms; only the agitation and depression subscale was reported.

Discussion

Summary

This review explored the behavioural and psychological impact of preferred music interventions on PWD. The results showed that diverse types of interventions used preferred

music as a medium across nine countries. This variation in interventions might be a reflection of the fact that the studies were conducted by researchers from different professional disciplines and published in different types of journals in order to target different audiences.

The diversity of interventions also meant that the studies varied in terms of music session duration and frequency. Studies often did not provide a clear rationale regarding the length, duration and frequency of music sessions. This finding is similar to that of Sung et al. (2005), who suggested that future studies could focus on identifying the optimum length and number of music sessions to provide guidance for the use of preferred music in clinical practice. The reviewed studies also did not report the time of the day and the season in which the intervention was carried out; these are thought to be potential variables that might affect PWD's mood (Sanchez et al., 2016). Furthermore, level of agitation (Gerdner, 2012), as well as, the accessibility of the intervention for those not living residential care were not heavily considered. Future studies could potentially investigate the best time for preferred music to be implemented. Lastly, as the majority of studies focused on older PWD in long-term care facilities, future studies could investigate earlier stages of dementia, as well as different types of dementia and whether the variation of settings, such as community centre or patient's home, may affect the outcomes.

Selecting preferred music

In terms of selecting the appropriate music for PWD, all studies took personal preferences into account. Most studies gathered this information by asking the PWD directly or asking their caregiver. Some of the studies used a validated music preference assessment tool when accessing PWD's personal preferences of music, such as the Assessment of Personal Music Preference (APMPQ) and The Music Preference Survey (MPS), whereas other studies gathered information through interviewing the PWD and caregiver. In one study, however, there was no report on the selected songs or type of music used for preferred

music interventions. Sung et al. (2005) suggested the possibility of using song lists and information gathered from the preferred music tools to create a database for older people in long-term care facilities. This list could be helpful for PWD living in care who no longer have close contact with their family caregivers, as well as those who are no longer able to verbalise their opinions. As a note of caution, however, clinicians need to be conscious of becoming over-reliant on a predetermined database as this is contrary to the purpose of creating an individualised playlist.

Study design

As this review excluded pilot studies, the sample sizes employed in the included studies were relatively larger than those reviewed by Sung et al. (2005), ranging from 22 to 6,298, as opposed to the three to 39 considered by the previous work. Larger sample size studies strengthens the external validity and reliability, suggesting that the findings might be more representative of the wider population of older PWD than the previous review. Furthermore, the review shows that since the publication of Sung et al. (2005), studies have used more rigorous research and experimental designs, such as RCTs, in testing the psychological effects of preferred music on PWD. Unfortunately, the review was unable to identify any qualitative studies that investigated the experiences of preferred music among those with a dementia. Qualitative methodologies could potentially help with gaining knowledge of PWD's experiences of preferred music, what preferred music means to them, and also their own interpretations of what such experiences mean (Atieno, 2009).

The MMAT (Hong et al., 2018) helped to further understand the methodological quality. The tool showed that some studies did not mention whether they had monitored participant adherence to the intervention. Low implementation fidelity might fail to distinguish the true effect of the intervention, or whether the effect stems from other extraneous variables (Carroll, Patterson, Wood, Booth, Rick & Balain, 2007). This also

meant that studies were prone to type III errors, whereby the lack significant results might not be caused by the intervention being ineffective itself, but rather the inadequacies in terms of participants adhering to the intervention (Carroll et al., 2007; Dobson & Cook, 1980). Secondly, the MMAT also helped to identify some studies that were unable to carry out blinding procedures, suggesting a potential risk of bias.

Outcome measures

The majority of studies used reliable and validated scales to measure psychological outcomes. Some of these tools were self-report measures, gathering information from participants and their caregivers. Although these tools are widely used, they are prone to socially desirable bias (Van De Mortel, 2008). Some studies employed observation measures. Observation methods are considered a gold standard for assessing PWD and they are helpful for gathering data from people who are no longer able to express themselves verbally (Curyto et al., 2008). Observation tools can risk experimenter bias unless interpreter reliability is established. Where feasible, blinding of the assessor also strengthens validity; however, this was not carried out in some of the reviewed studies.

Sakamoto et al. (2013) used the facial scale as an observational tool to evaluate PWD's changes in their emotions and they appeared to assume that facial changes represented emotional changes. It is worth noting, however, that some PWD experience a loss of facial expression, including their ability to smile, depending on the type of dementia. This is particularly evident in Lewy body dementia (McKeith, 2004), which therefore may potentially have been misinterpreted as an indicator of the presence of pain, although it could reflect other emotions and issues as well (Sheu, Versloot, Nader, Kerr & Craig, 2011). Other neurological comorbidities, such as the pseudobulbar affect, which is present in 10 to 38.5 percent of dementia diagnoses, might also interfere with observation of facial expression (Colamonico, Formella & Bradley, 2012; Work, Colamonico, Bradley & Kaye, 2011). These

factors could affect the reliability of using facial expression as a way to understand emotions. Sakamoto et al. (2013) used heart rate as a measure to find PWD's emotional state; the advantage of using this physiological measure is that it reduces the likelihood of the participant and assessor contributing biases. However, measuring heart rate can be somewhat indirect and further research is needed to investigate whether heart rate and overall cardiac pattern is a good indicator of emotional change.

Findings

The review found inconsistent results on the effect of preferred music interventions in reducing agitation. This finding is contrary to the previous reviews, which suggested that preferred music has a positive outcome in reducing agitation. There is a possibility that this difference is a result of having dissimilar inclusion criteria, where pilot studies that adopted small sample sizes and employed less rigorous research designs were previously included (Sung et al., 2005; Vink et al., 2003). This inconsistency was found despite the majority of the studies using the same tool to measure agitated behaviours, which raises questions about the generalisability of the study findings presented in the review. Further investigation is needed to investigate whether any confounding variables might have contributed to differences in the results across the studies.

In terms of the effect of preferred music intervention on level of anxiety, the majority of studies used the same outcome measure. Despite some studies showing that preferred music intervention performed better than control conditions (standard care and social activities), there was little evidence to suggest that preferred music reduced PWD anxiety over time. Inconsistent results were also found in terms of depression. The fact that the studies used different tools to measure changes in PWD's level of depression made it harder to draw definitive conclusions. These findings are in line with other previous reviews on non-preferred music interventions where evidence was found to be inconclusive (Brotons et al.,

1997; McDermott, Crellin, Ridder & Orrell, 2012; Koger, Chapin & Brotons, 1999; Nilsson, 2008; Vink et al., 2003, 2011).

Inconsistent results were also found on the effect of preferred music on the overall emotional state of those with dementia. Of the three studies that investigated overall emotional state, one of the studies did not report detailed results. The two studies that provided more detail used different tools but showed inconsistencies in their results, suggesting that further investigation is needed. The use of evaluation tools that look at facial changes and heart rate raises questions about whether either measure directly reflects PWD's internal emotional state (e.g. Thomas, Crutch & Camic, 2017). Lastly, the review found a reduction in medication use after receiving preferred music. This suggests that preferred music may be an effective intervention to manage behavioural and psychological symptoms.

Limitations.

It is worth noting that this current review has its own limitations. Firstly, it closely adopted Sung & Marie's (2005) review, to offer a more up to date understanding of preferred music's effect. However, it later came to the author's attention that the search terms might potentially be incomplete. For example, Sung & Marie's (2005) did not capture American spelling such as the word "individualized". Furthermore, root words could also benefit by using truncation (e.g. favour*) to broaden the search. This meant that some publications might have been missed. On the other hand, despite this present review having adopted Sung & Marie's (2005) search strategies, the databases used in the present review have a "relevant term" functions, which reduced the likelihood that relevant papers were missed. Nonetheless, a future review could benefit from revising the search strategies to ensure that relevant publication are not missed.

Research recommendations

Future research can strengthen and build on the foundations established by these studies in a number of ways. Further assurance could be achieved by using double blinding in RCTs, adhering to treatment fidelity to avoid potential type III errors, and using consistent investigative tools for the constructs of agitation, anxiety, depression and overall emotional state. For agitation, additional consideration of the time of day when agitative behaviours occur would add further precision to the measurement and assessment of the impact of the intervention. Future research could also investigate other potential factors, such as treatment and follow-up duration (Sung et al., 2010), optimum time for conducting the intervention (Sakamoto et al., 2013), type of songs (Guetin et al., 2009), cognitive function (Cheung et al., 2018) and controlling usual care conditions (Cheung et al., 2018). Ledger and Baker (2007) also suggested the importance of future studies accounting for outside factors that might influence the PWD, including illnesses, hospitalisations, medication, environmental changes, interventions received and deaths among family and friends.

Qualitative studies were absent from this area and thus they could be used to add further information about the nuances of the experience of PWD, caregivers and staff. Specifically, interviewing nursing staff in long-term care facilities to assess their attitudes and perceptions of the implementation of preferred music interventions would be helpful. By using mixed methodologies future research can also develop easy-to-use observation tools to examine the effect of preferred music across different levels of impairment and diagnosis type in community, day and residential care settings. Future studies could also consider investigating the effect of preferred music on less traditional outcomes and "non-symptom" related outcomes, such as engagement and social interactions, which have been shown to promote the wellbeing of PWD (Mitchell & Agnelli, 2015).

As discussed previously, the intentional focus of the review was to investigate the effects of preferred music alone to offer a more detail description of the intervention. A future review might benefit from comparing the effects between preferred music and non-preferred music, reviewing comparison studies of the two interventions, as well as comparing their effect sizes across studies.

Clinical implications

There are several clinical implications that clinical psychologists and others working in older adult care can *cautiously* consider as a result of this review. Firstly, as a way to support personhood (Kitwood, 1997), using a quantitative tool to assist selection of preferred music could benefit PWD with mild to moderate impairments (Unadkat et al., 2017) to ensure music that is important to them is selected. Secondly, creating a song database for people who are unable to verbally express themselves would allow for greater national access to a range of music across the types of dementia and it would be useful for family members, paid carers and healthcare professionals. Thirdly, it is possible to be moderately confident that for some PWD preferred music should be considered on at least a trial basis as an alternative way of managing BPSD. Lastly, the inconsistent findings in the review raise questions about the effectiveness of using traditional symptom-based measures in clinical settings and suggest the need to find other more effective measures of the impact of interventions on the wellbeing of PWD.

Conclusions

The results of this review indicate that the findings on the effect preferred music has on PWD are inconsistent. These inconsistent results were found across all outcomes reviewed, including agitation, anxiety, depression and overall emotional state. This finding is also contrary to previous reviews that studied researches using a smaller sample size. Having such

mixed results meant that it was difficult for the current review to draw any meaningful conclusions. Furthermore, the review identified a number of methodological issues that might have caused biases in the studies and affected the accuracy of the results.

Future studies investigating this area could consider adopting a more vigorous design in addressing biases that could affect the investigation of the true effect of preferred music on PWD. Nonetheless, clinicians could cautiously consider the use of preferred music for PWD, as well as considering using alternative, "non-symptom" related outcomes to measure intervention effectiveness.

References

Akobeng, A. K. (2005). Understanding randomised controlled trials. *Archives of Disease in Childhood*, *90*, 840-844. doi: 10.1136/adc.2004.058222

Alzheimer's Research UK, (2013, Dec 7). *135 million people will live with dementia worldwide by 2050: Alzheimer's Disease International announces new global prevalence gold figures*. Retrieved from https://www.dementiastatistics.org/statistics-about-dementia/

Alzheimer's Society. (2014). *Dementia UK: Second edition – Overview*. London, UK: Alzheimer's Society. Retrieved from http://eprints.lse.ac.uk/59437/1/Dementia_UK_Second_edition_-_Overview.pdf

Alzheimer's Society. (2018). *The dementia guide: Living well after diagnosis*. London: Alzheimer's Society.

Atieno, O. P. (2009). An analysis of the strengths and limitation of qualitative and quantitative research paradigms. *Problems of Education in the 21st Century*, *13*, 13-38. Retrieved from http://www.scientiasocialis.lt/pec/node/files/pdf/Atieno_Vol.13.pdf

Bidewell, J. W., & Chang, E. (2011). Managing dementia agitation in residential aged care. *Dementia*, *10*, 299-315. doi: 10.1177/1471301211407789

Bradt, J., Burns, D. S., & Creswell, J. W. (2013). Mixed methods research in music therapy research. *Journal of Music Therapy*, *50*, 123-148. doi: 10.1093/jmt/50.2.123

Brotons, M., Koger, S. M., & Pickett-Cooper, P. (1997). Music and dementias: A review of literature. *Journal of Music Therapy*, *34*, 204-245. doi: 10.1093/jmt/37.3.183

Buckwalter, K. C., & Hall, G. (1987). Families of the institutionalized older adult: A neglected resource. In T. H. Brubaker (Ed.), *Aging, health and family: Long term care* (pp. 176-196). Beverly Hills: Sage.

Burnside, I. (1988). *Dementia and delirium.* In I. Burnside (Ed.), *Nursing and the aged* (pp. 732-794). New York McGraw-Hill.

Camic, P. M., Williams, C. M. & Meeten, F. (2011: on-line, 2013: print). Does a 'singing together group' improve the quality of life of people with a dementia and their carers? A pilot evaluation study. *Dementia*, *12*, 152-171. doi: 10.1177/1471301211422761

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, *2*, 40-49. doi: 10.1186/1748-5908-2-40

Cheung, D. S. K., Lai, C. K. Y., Wong, F. K. Y., & Leung, M. C. P. (2018). The effects of the music-with-movement intervention on the cognitive functions of people with moderate dementia: A randomized controlled trial. *Aging & Mental Health*, *22*, 306-315. doi: 10.1080/13607863.2016.1251571

Chlan, L. (1998). Effectiveness of a music therapy intervention on relaxation and anxiety for patients receiving ventilatory assistance. *Heart & Lung: The Journal of Acute and Critical Care*, *27*, 169-176. doi: 10.1016/S0147-9563(98)90004-8

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. doi: 10.1111/j.1532-5415.1986.tb04303.x

Cohen-Mansfield, J., Dakheel-Ali, M., & Marx, M. S. (2009). Engagement in persons with dementia: the concept and its measurement. *The American Journal of Geriatric Psychiatry*, *17*, 299-307. doi: 10.1097/JGP.0b013e31818f3a52

Colamonico, J., Formella, A., & Bradley, W. (2012). Pseudobulbar affect: burden of illness in the USA. *Advances in Therapy*, *29*, 775-798. doi: 10.1007/s12325-012-0043-7

Cooke, M. L., Moyle, W., Shum, D. H., Harrison, S. D., & Murfield, J. E. (2010). A randomized controlled trial exploring the effect of music on agitated behaviours and anxiety in older people with dementia. *Aging and Mental Health*, *14*, 905-916. doi: 10.1080/13607861003713190

Curyto, K. J., Van Haitsma, K., & Vriesman, D. K. (2008). Direct observation of behavior: A review of current measures for use with older adults with dementia. *Research in Gerontological Nursing*, *1*, 52-76. doi: 10.3928/19404921-20080101-02

Davis, W., & Hadley, S. (2015). A history of music therapy. *Music therapy handbook*, 17-28.

Dobson, D., & Cook, T. J. (1980). Avoiding type III error in program evaluation: Results from a field experiment. *Evaluation and Program Planning*, *3*, 269-276. doi: 10.1016/0149-7189(80)90042-7

Evans, D. (2002). The effectiveness of music as an intervention for hospital patients: A systematic review. *Journal of Advanced Nursing*, 37, 8-18. doi:10.1046/j.1365- 2648.2002.02052.

Finkel, S. I., Lyons, J. S., & Anderson, R. L. (1992). Reliability and validity of the Cohen–Mansfield agitation inventory in institutionalized elderly. *International Journal of Geriatric Psychiatry*, *7*, 487-490. doi: 10.1002/gps.930070706Kid

Garland, K., Beer, E., Eppingstall, B., & O'Connor, D. W. (2007). A comparison of two treatments of agitated behavior in nursing home residents with dementia: simulated family presence and preferred music. *The American Journal of Geriatric Psychiatry*, *15*, 514-521. doi: 10.1097/01.JGP.0000249388.37080.b4

Gerdner, L. A. (1992). *The effects of individualized music on elderly patients who are confused and agitated* (Doctoral dissertation, University of Iowa).

Gerdner, L. (1997). An individualized music intervention for agitation. *Journal of the American Psychiatric Nurses Association, 3*, 177-184. doi: 10.1016/S1078-3903(97)90043-4

Gerdner, L. A. (2000). Effects of individualized versus classical “relaxation” music on the frequency of agitation in elderly persons with Alzheimer's disease and related disorders. *International Psychogeriatrics*, *12*, 49-65. doi: 10.1017S1041610200006190

Gerdner, L. A. (2012). Individualized music for dementia: Evolution and application of evidence-based protocol. *World Journal of Psychiatry*, *2*, 26-32. doi: 10.5498/wjp.v2.i2.26

Guetin, S., Portet, F., Picot, M. C., Pommié, C., Messaoudi, M., Djabelkir, L., ... & Touchon, J. (2009). Effect of music therapy on anxiety and depression in patients with Alzheimer’s type dementia: Randomised, controlled study. *Dementia and Geriatric Cognitive Disorders*, *28*, 36-46. doi: 10.1159/000229024

Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., ... & O’Cathain, A. (2018). Mixed methods appraisal tool (MMAT), version 2018. *IC Canadian Intellectual Property Office, Canada*. Retrieve from http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf

Kitwood, T. M. (1997). *Dementia reconsidered: The person comes first*. Buckingham, UK: Open university press.

Koger, S. M., Chapin, K., & Brotons, M. (1999). Is music therapy an effective intervention for dementia? A meta-analytic review of literature. *Journal of Music Therapy*, *36*, 2-15. doi: 10.1093/jmt/36.1.2

Kuring, J. K., Mathias, J. L., & Ward, L. (2018). Prevalence of depression, anxiety and PTSD in people with dementia: A systematic review and meta-analysis. *Neuropsychology Review*, *28*, 393-416. doi: 10.1007/s11065-018-9396-2

Kwak, J., Anderson, K., & O'Connell Valuch, K. (2018). Findings from a prospective randomized controlled trial of an individualized music listening program for persons with dementia. *Journal of Applied Gerontology, 1-16*. doi: 10.1177/0733464818778991.

Ledger, A. J., & Baker, F. A. (2007). An investigation of long-term effects of group music therapy on agitation levels of people with Alzheimer's Disease. *Aging and Mental Health*, *11*, 330-338. doi: 10.1080/13607860600963406

Leon, A. C., Davis, L. L., & Kraemer, H. C. (2011). The role and interpretation of pilot studies in clinical research. *Journal of Psychiatric Research*, *45*, 626-629. doi: 10.1016/j.jpsychires.2010.10.008

Lou, M. F. (2001). The use of music to decrease agitated behaviour of the demented elderly: The state of the science. *Scandinavian Journal of Caring Sciences*, *15,* 165-173. doi: 10.1046/j.1471-6712.2001.00021.x

McDermott, O., Crellin, N., Ridder, H. M., & Orrell, M. (2012). Music therapy in dementia: A narrative synthesis systematic review. *International Journal of Geriatric Psychiatry*, *28*, 781-794. doi: 10.1002/gps.3895

McKeith, I. (2004). Dementia with Lewy bodies. *Dialogues Clinical Neuroscience*, *6*, 333-341. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3181810/pdf/DialoguesClinNeurosci-6-333.pdf

Mitchell, G., & Agnelli, J. (2015). Person-centred care for people with dementia: Kitwood reconsidered. *Nursing Standard, 30*, 46. doi: 10.7748/ns.30.7.46.s47

Nilsson, U. (2008). The anxiety-and pain-reducing effects of music interventions: a systematic review. *AORN Journal*, *87*, 780-807. doi: 10.1016/j.aorn.2007.09.013

Nolan M, Brown J, Davies S, Nolan J, Keady J (2006). *The Senses Framework: improving care for older people through a relationship-centred approach. Getting Research into Practice (GRiP) Report No. 2.* Sheffield: Sheffield Hallam University. Retrieved from http://shura.shu.ac.uk/280/1/pdf_senses_framework_report.pdf

Orgeta, V., Qazi, A., Spector, A., & Orrell, M. (2015). Psychological treatments for depression and anxiety in dementia and mild cognitive impairment: systematic review and meta-analysis. *The British Journal of Psychiatry*, 207, 293-298. doi: 10.1192/bjp.bp.114.148130

Palm, R., Sorg, C. G., Ströbel, A., Gerritsen, D. L., & Holle, B. (2018). Severe Agitation in Dementia: An Explorative Secondary Data Analysis on the Prevalence and Associated Factors in Nursing Home Residents. *Journal of Alzheimer's Disease*, *66,* 1463-1470. doi: 10.3233/JAD-180647

Pedersen, S. K., Andersen, P. N., Lugo, R. G., Andreassen, M., & Sütterlin, S. (2017). Effects of music on agitation in dementia: A meta-analysis. *Frontiers in psychology*, 8, 742. doi: 10.3389/fpsyg.2017.00742

Perlman, C. M., & Hirdes, J. P. (2008). The aggressive behavior scale: A new scale to measure aggression based on the minimum data set. *Journal of the American Geriatrics Society*, *56*, 2298-2303. doi: 10.1111/j.1532-5415.2008.02048.x

Pluye, P., Robert, E., Cargo, M., Bartlett, G., O'cathain, A., Griffiths, F., ... & Rousseau, M. C. (2011). *Proposal: A mixed methods appraisal tool for systematic mixed studies reviews*. Montréal: McGill University, *2*, 1-8. Retrieved from http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/84371689/MMAT%202011%20criteria%20and%20tutorial%202011-06-29updated2014.08.21.pdf

Raglio, A., Bellandi, D., Baiardi, P., Gianotti, M., Ubezio, M. C., Zanacchi, E., ... & Stramba-Badiale, M. (2015). Effect of active music therapy and individualized listening to music on dementia: a multicenter randomized controlled trial. *Journal of the American Geriatrics Society*, *63*, 1534-1539. doi: 10.1111/jgs.13558

Ragneskog, H., Asplund, K., Kihlgren, M., & Norberg, A. (2001). Individualized music played for agitated patients with dementia: Analysis of video-recorded sessions. *International Journal of Nursing Practice*, *7*, 146-155. doi: 10.1046/j.1440-172X.2001.00254.x

Randall, T. (1991). Music not only has charms to soothe, but also to aid elderly in coping with various disabilities. *JAMA*, *266*, 1323-1329. doi:10.1001/jama.1991.03470100015005

Rare Dementia Support (n.d.) *Specialist support for people affected by one of five rare dementia diagnoses*. Retrieved from http://www.raredementiasupport.org/

Robb, S. L., Hanson-Abromeit, D., May, L., Hernandez-Ruiz, E., Allison, M., Beloat, A., ... & Polasik, S. (2018). Reporting quality of music intervention research in healthcare: A systematic review. *Complementary therapies in medicine*, 38, 24-41. doi: 10.1016/j.ctim.2018.02.008

Sakamoto, M., Ando, H., & Tsutou, A. (2013). Comparing the effects of different individualized music interventions for elderly individuals with severe dementia. *International Psychogeriatrics*, *25*, 775-784. doi: 10.1017/S1041610212002256

Sánchez, A., Maseda, A., Marante-Moar, M. P., de Labra, C., Lorenzo-López, L., & Millán-Calenti, J. C. (2016). Comparing the effects of multisensory stimulation and individualized music sessions on elderly people with severe dementia: A randomized controlled trial. *Journal of Alzheimer's Disease*, *52*, 303-315. Retreived from https://ruc.udc.es/dspace/bitstream/handle/2183/16601/S%C3%A1nchez_ComparingEffects.pdf?sequence=2&isAllowed=y

Shankar, K. K., Walker, M., Frost, D., & Orrell, M. W. (1999). The development of a valid and reliable scale for rating anxiety in dementia (RAID). *Aging & Mental Health*, *3*, 39-49. doi: 10.1080/13607869956424

Sheu, E., Versloot, J., Nader, R., Kerr, D., & Craig, K. D. (2011). Pain in the elderly: Validity of facial expression components of observational measures. *The Clinical Journal of Pain*, *27*, 593-601. doi: 10.1097/AJP.0b013e31820f52e1

Sung, H. C., & Chang, A. M. (2005). Use of preferred music to decrease agitated behaviours in older people with dementia: A review of the literature. *Journal of Clinical Nursing*, *14*, 1133-1140. doi: 10.1111/j.1365-2702.2005.01218.x

Sung, H. C., Chang, A. M., & Abbey, J. (2006). The effects of preferred music on agitation of older people with dementia in Taiwan. *International Journal of Geriatric Psychiatry: A Journal of The Psychiatry of Late Life and Allied Sciences*, *21*, 999-1000. doi: 10.1002/gps.1585

Sung, H. C., Chang, A. M., & Lee, W. L. (2010). A preferred music listening intervention to reduce anxiety in older adults with dementia in nursing homes. *Journal of Clinical Nursing*, *19*, 1056-1064. doi: 10.1111/j.1365-2702.2009.03016.x

Thiese, M. S. (2014). Observational and interventional study design types; An overview. *Biochemia Medica, 24*, 199-210. doi: 10.11613/BM.2014.022

Thomas, K. S., Baier, R., Kosar, C., Ogarek, J., Trepman, A., & Mor, V. (2017). Individualized music program is associated with improved outcomes for US nursing home residents with dementia. *The American Journal of Geriatric Psychiatry*, *25*, 931-938. doi: 10.1016/j.jagp.2017.04.008

Thomas, G. E. C., Crutch, S. J. & Camic, P. M. (2017). Measuring physiological responses to the arts in people with a dementia: A systemized literature review. *International Journal of Psychophysiology, 123,* 64-73. doi:10.1016/j.ijpsycho.2017.11.008

Unadkat, S., Camic P.M. & Vella-Burrows, T. (2017). Understanding the experience of group singing for couples where one partner has a diagnosis of dementia. *The Gerontologist, 47*, 468-478. doi:10.1093/geront/gnv698

Van de Mortel, T. F. (2008). Faking it: Social desirability response bias in self-report research. The *Australian Journal of Advanced Nursing, 25*, 40-48. Retrieved from https://epubs.scu.edu.au/cgi/viewcontent.cgi?article=1001&context=hahs_pubs

Vink, A. C., Bruinsma, M. S., & Scholten, R. J. (2003). Music therapy for people with dementia. *Cochrane database of systematic reviews*, 4. Retrieved from https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD003477.pub2/abstract

Work, S. S., Colamonico, J. A., Bradley, W. G., & Kaye, R. E. (2011). Pseudobulbar affect: an under-recognized and under-treated neurological disorder. *Advances in Therapy*, *28*, 586-601. doi: 10.1007/s12325-011-0031-3

World Health Organization. (2018). *Towards a dementia plan: A WHO guide*. Geneva, Switzerland: World Health Organization. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/272642/9789241514132-eng.pdf

Zare, M., Birashk, B., & Ebrahimi, A. A. (2012). Reliability, validity and factor analysis of Cohen-Mansfield Agitation Inventory (CMAI). *Iranian Journal of Psychiatry & Clinical Psychology*, *18, 67-73*. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&AuthType=cookie,ip,shib&db=asx&AN=79330382&site=eds-live

Section B: Journal Article

DEVELOPMENT OF THE VIDEO ANALYSIS SCALE OF ENGAGEMENT (VASE) TOOL FOR PEOPLE WITH ADVANCED DEMENTIA

Word Count 7998 (81)

A thesis submitted in partial fulfilment of the requirements of
Canterbury Christ Church University for the degree of
Doctor of Clinical Psychology

APRIL 2019

SALOMONS
CANTERBURY CHRIST CHURCH UNIVERSITY

Abstract

The current study sought to develop a valid, reliable and unobtrusive tablet computer-based observational tool to appraise a continuous scale of engagement with people with advanced dementia. VASE was designed to enable the rating of moment-by-moment changes in engagement during an intervention, which would be useful for process evaluation in research. An initial version of the Video Analysis Scale of Engagement (VASE) was tested. Face validity and content validity were conducted to validate an operational definition of engagement and develop an acceptable protocol for the tool. Thirty-seven non-professional and professional volunteers were recruited to view and rate people with dementia's level of engagement in the music activities using the VASE. An inter-class coefficient (ICC) test gave a high level of rating agreement across professionals and non-professionals. However, the ICC results of within-professionals were mixed. Mixed-linear modelling suggested there that the types of interventions (active or passive music listening), the particular intervention session being rated, five second "stages" of each video and the age of those doing the rating could affect the ratings. Results suggested that raters used the VASE in a dynamic fashion, and that the tool was able to distinguish between interventions. Further investigation and adjustments are warranted for this to be considered a valid and reliable tool in the measurement of engagement of people with advanced dementia in a group activity setting.

Introduction

Dementia is a growing challenge for the UK and has been set as one of the priorities for NHS England and the government (Alzheimer's Society, 2015). In 2015 the Prime Minister launched the "Challenge on Dementia 2020" (Department of Health, 2015) with the vision that the UK would become one of the leading countries in the world for dementia care and research. Non-pharmacological interventions (NPI) are thought to be a meaningful and preferred treatment of care for PWD (Cabrera et al., 2015). These interventions include, for instance, cognitive stimulation therapy (Orrell et al., 2017), music and art therapies (Raglio et al., 2015; Deshmukh, Holmes & Cardno, 2018), talking therapies (Cheston & Ivanecka, 2017), and others. The evaluation of whether these interventions are successful often relies on self-report measures or staff observations of behavioural and psychological symptom changes. Although often become increasingly socially isolated due to their progressive deterioration in communication abilities, regrettably social interaction and engagement have not been considered as useful outcomes in dementia care (Sung & Chang, 2005). More recently, however, there has been an emphasis on promoting welling for PWD through social engagement within their immediate environments (Martyr, et al., 2018). Such an approach recognizes that focusing solely on emotional and behavioural outcomes to determine if an intervention is successful might pose a danger where interventions are "prescribed" to PWD based on those outcomes, without considering the individual's personal choice.

Beard, Knauss, and Moyer (2009) found that people with dementia (PWD) continued to want an enriched life after being diagnosed. Having positive attitudes, and engaging with physical, mental and social activities are thought to be helpful to maintain an enriched life (Beard et al., 2009). Kitwood's (1997) person-centered approach highlighted the importance of maintaining personhood in dementia care, where it stresses the need promoting social

inclusion (e.g., stay within social circles, during conversations) to maintain identity for the individual. The maintenance of inclusion and identity could be achieved through self-directed support in dementia care (Mental Health Foundation, 2011), such as placing a strong emphasis on the importance to respond to the needs "in the moment" (Dunne, 2002).

Understanding the choices and needs of people at advanced stages of dementia is not always clear as deterioration in memory, difficulties in communication and impairments in daily activities increase as the disease progresses. Consequently, being able to express basic needs and wants becomes more difficult and PWD voices become "lost" (Lai, 2015; Kumar & Kuriakose, 2013). Providing appropriate care that is meaningful for the PWD relies on staff and carers' understanding and familiarity with the individual, which is enhanced by careful observation of daily interactions (Mental Health Foundation, 2011).

Engagement

One way to assist staff and carers to understand PWD's preferences is to consider PWD's level of engagement with particular activities. Engagement, as a form of social interaction, is thought to be an essential aspect in determining the effectiveness of interventions in their ability to promote meaningful activity (Jones, Sung & Moyle, 2015). One of the key aspects of person-centred care is the recognition that all human life is grounded in social relationships (Brooker, 2015). Therefore, advanced dementia care should focus on creating a rich social environment to foster personal growth by maintaining engagement, relationships and activities appropriate to the level of impairment (Mitchell & Agnelli, 2015). Other than human interaction, Cohen-Mansfield, Dakheel-Ali and Marx (2009) suggested that engagement could refer to "the act of being occupied or involved with an external stimulus" (p. 2); this idea suggests that being engaged not only means connecting with people, but also with other stimuli, such as objects, music and activities.

Engagement in the social context is often referred to an individual's participation in the activities of a social group (Prohaska et al., 2012). Zhang, Jiang, and Carroll (2011) suggested that engagement means that the member stays in the group and interacts with others. Perugia et al. (2018) defined engagement as wellbeing, enjoyment and active involvement triggered by meaningful activities in PWD. For the purpose of this study, engagement is conceptualized as a state of wellbeing and involvement triggered by activities in a group.

During the advanced stages of dementia, when language skills often deteriorate, PWD frequently reside in residential care (Herrmann & Gauthier, 2008). This can create difficulties in day-to-day social life and as a result engagement in daily activities diminishes over time (Claire, Mathews & Kosloski, 2005; Hubbart, Cook, Tester & Downs, 2002). Reduced engagement can lead to boredom, loneliness and depression (Cohen-Mansfield et al., 2009); it is therefore essential for PWD to participate in activities that promote positive social interactions in residential care. Research further suggests that engaging in meaningful and enjoyable activities can lead to a better quality of life (Smit, Lange, Williemse, Twisk & Pot, 2016), fewer behaviour problems (Braun, 2019) and increased positive emotions (Kolanowski, Van Haitsma, Meeks & Litaker, 2014).

It is important to note that a lower level of engagement in particular activities should not be seen as a symptom or a lack of ability, but rather, may be an indicator of having a strong sense of self (Sabat, 2006). Individuals rejecting participation in an activity might be indicating an ability to advocate and express needs. Sabat (2006) proposed that "self" remains even in the advanced stages of dementia and there are three forms of self, each with different attributions. The most vulnerable self for people in the advanced stage of dementia is a "publicly presented persona that requires the cooperation of others". To protect this part

of the self, Sabat (2006) stressed that carers should provide good quality interactions that support relationships and the role of the individual. Sabat's (2006) theory is a further development of Kitwood's concept of personhood focusing on person-centred care, which identifies activities to stimulate engagement at advanced stages of dementia. As such, gauging engagement in PWD becomes an important issue.

Methods to measure engagement

A review conducted by Curyto, Van Haitsma and Vriesman (2008) concluded that observation tools are the gold standard in assessment for older people at the advanced stages of dementia. Observation techniques have previously been used to investigate the process and interactions in dementia care (Curyto et al., 2008; Engström, Marmstål, Hammar, Williams & Götell, 2011; Gaugler, Hobday & Savik, 2013). Observations are thought to be able to gather meaningful data about the functioning of PWD that might be missed by standard questionnaires (Algar, Woods & Windle, 2014).

Limitations to the traditional observational tools include requiring observers to be present during a given session, which is costly and labour-intensive, as well as potentially creating stress that could affect group interactions (Carthey, 2003). Less intrusive and cost-effective alternatives could address these concerns. Robert et al. (2010) noted that the existing engagement observational measures are difficult to use, even for professionals. A previous review considered 68 available assessment scales for Alzheimer's disease for various purposes; it was concluded that there was a need for an "easy-to-administer" scale for identifying response to therapy in daily practice (Robert, Ferris, Gauthier, Ihl, Winblad & Tennigkeit, 2010). The existing observational protocols are time-intensive to administer and frequently unable to monitor direct therapist-PWD interactions that reflects person-centred care (Gaugler et al., 2013). For example, the Observational Measurement of Engagement

(OME) has been validated to examine the engagement of PWD in interventions such as music (Cohen-Mansfied et al., 2009). This follows a complex protocol requiring formal training and a substantial amount of time to learn. Its complexity potentially reduces the accessibility of such measures for non-researchers. Consequently, staff carers might struggle to find appropriate tools to assess whether people at the advanced stages of dementia are engaged in particular activities that they preferred. Furthermore, the available engagement measures constitute either a single score system or an average score from a certain period (e.g. 30 seconds to 1 minute). However, human interactions are "dynamic", changing "moment-by-moment", which this method of measuring might miss. Therefore, there is a need create a flexible tool that allows those doing the rating (henceforth "raters") to capture these dynamic changes as and when they observe them.

Video-based observation

Video-based observation offers a good alternative as it was developed to be an unobtrusive method, minimising disruption to the social setting through the presence of researchers and observers. This method can also capture multiple, complex interactions simultaneously while gathering a larger amount of data than traditional observational methods (Asan & Montague, 2014). A further benefit of using video analysis with a severely impaired dementia population is that it allows both researchers and care staff to closely view group interactions, meanwhile facilitating an examination and understanding of subtle behaviours occurring within the group (Asan et al., 2014). Video analysis can also enable the use of raters from the wider social system (e.g. lay people, family carers and non-dementia experts), therefore permiting a more comprehensive perspective of care. A wider perspective of care, including families views and support, is thought to be important in dementia care (Moore et al., 2014).

In order to fully utilise the potential of video-based observation, this current project aims to design a more straightforward measure of engagement for use with video recorded observation. This tool will also incorporate a time-tracking system, enabling the raters to capture moment-by-moment changes in the video and to rate the level of engagement as they observe those in care. The measure will be developed to observe the level of engagement in people with advanced dementia in response to music activities.

Music interventions as intervention

The rationale behind using music activities as a way of testing the observation tool developed is based on the fact that music intervention has been recommended as an appropriate intervention for PWD (Alzheimer's Society, 2018; Abraha et al., 2017). Music interventions have also been heavily promoted in dementia settings and care homes (e.g. individualised music, and "A Choir in Every Care Home" initiative). These interventions have been found to increase PWD's quality of life (QoL) (Vasionytė & Madison, 2013)., as well as promoting other physio-psychosocial outcomes in dementia (Cooke, Moyle, Shum, Harrison & Murfield, 2010). This physical and psychological improvement echoes Murrock and Higgin's (2009) theory of music, mood and movement. Most importantly, such interventions have been found to improve levels of engagement (Eggert et al., 2015), and is therefore a suitable non-pharmacological intervention for the purpose of this study. Music interventions are thought to engage the PWD by promoting relaxation or creating a sensory stimulant (Cohen-Mansfield et al., 2009).

Music intervention is of particular importance for people at the advanced stage of dementia, as even though their language ability has deteriorated, their musical abilities are relatively preserved (Baird & Samson, 2015). This finding supports the theory of individualised music intervention for agitation (IMIA), which suggests that music acts as a

medium for communication for PWD with an impaired ability to understand verbal language (Gerdner, 2012). Currently, the one music program available in the UK for advanced dementia is the Music for Life (MfL) program (Music for Life, 2014), which is what was chosen to be activity to test out the VASE. MfL is an approach that aims to bring together professional musicians, care staff and people living with dementia through interactive music to enhance their quality of life (Rose, 1993).

Aim and Objectives

An effective and easy-to-administer tool that assesses engagement in advanced dementia does not yet exist so far as the researcher is aware. The present study aimed to develop a non-intrusive user-friendly video-based observational tool to assess level of engagement during a music intervention (MfL): the Video Analysis Scale of Engagement (VASE). The study was carried out in two stages: the protocol development stage and the validation stage. This paper presents the development of the VASE and reports its psychometric properties in terms of validity and reliability analyses. The research question, objectives and hypotheses of the study are presented in Table 1.

Table 1. *Project objectives and hypotheses*

Primary research question

Can an observational rating tool effectively measure the engagement of people with advanced dementia in music activity?

Overall objectives

To develop an observational rating tool with a user-friendly operational protocol to measure the level of engagement of people with advanced dementia. The intervention used for observing the participants' level of engagement is a music-based activity (Music for Life) in residential care.

Objective 1. Identify an operational definition of engagement for the VASE and determine its face and content validity.

Objective 2. Determine inter-rater reliability and intra-rater reliability of VASE by comparing groups of raters (non-professionals and dementia care professionals). If the scale is found to be reliable, assess the validity of the scale to differentiate engagement by comparing two conditions, namely the Music for Life group (MfL) and passive listening (PL) groups).

Objective 3. Assess whether other variables, including age, gender, ability to play a musical instrument, participation in a choir in the past or the present, types of session (MfL versus PL) or the order of videos raters rated (order effect), might affect the engagement score.

Hypotheses

Hypothesis 1. Related to objective 2 above, there will be a significant correlation between the VASE rating by dementia experts and non-professionals.

Hypothesis 2. Related to objective 2, there will be a correlation between the VASE rating among the dementia experts as well as that among the non-professionals.

Hypothesis 3. Related to objective 2, the rating tool will be sufficiently sensitive to differentiate the engagement level between the two music-based activities (MfL versus PL).

Hypothesis 4. Related to objective 3, extraneous variables will not affect the rating.

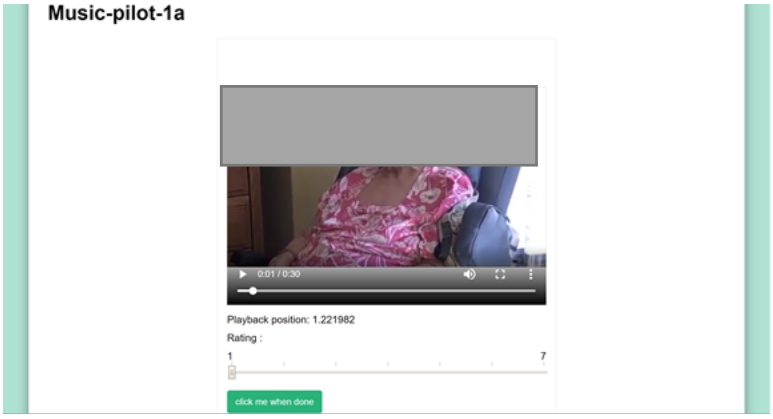
Methods

Design and procedure

This study adopted a mixed-methods design study (Creswell, Plano Clark, Gutmann & Hanson, 2003) and was separated into two stages: the development stage and the experiment stage. In the development stage, a qualitative method was used to identify an operational definition of engagement, followed by developing a protocol for the VASE. Face validity and content validity were then assessed. Face validity is the appropriateness, sensibility, or relevance of the tool and its items as they appear to the persons answering the test (Holden, 2010), whereas content validity investigates whether tool adequately covers the content that it should be, with respect to the variable (i.e. engagement) (Heale & Twycross, 2015). The experimental stage adopted a quantitative approach that followed a quasi-experimental design to investigate the detection and rating of different levels of engagement in an active MfL group and a passive PL group. The reliability of the VASE was then assessed. The following section first describes the VASE scale and then discusses the steps entailed in the two different stages. Other validity such as criterion validity was considered to compare how well the VASE is in measuring engagement against another validated tool (Heale & Twycross, 2015). This process would involve participants rating both tools. This validity test was not carried out due to two reason, firstly, there were no tool available that measure moment-by-moment level of engagement. Secondly, if the VASE was to compare with other "non" moment-by-moment engagement tools, such as OME; participants would have to pause the VASE rating every 15 second to complete other rating, making the duration of the experiment much longer, more complex and more difficult for participants focus on watching and rating the video itself.

Measures

Video analysis scale of engagement (VASE). The VASE (Figure 1) is an offline application written in the hypertext markup language (HTML). It was developed by the Created out of Mind research team (Appendix A), established in 2016 by a diverse team of collaborators in various professions and capacities. The VASE is a computer tablet-based program. The program consists of three main parts: 1) a 500 mm x 400 mm viewing box that allows raters to view a preloaded video; 2) a 7-point scale for raters to record changes in engagement while viewing the video; and 3) an exact time stamp to determine time of rating (Figure 1). The tool adopts a continuous scoring system using a 7-point Likert-type scale to assess level of engagement whilst viewing a video segment. Responses and time of the responses are automatically recorded by the software. The VASE enables raters to respond in real time by simply tapping a scale on the screen without the need to stop the footage to record ratings. In an earlier version of the VASE by the research team, raters were asked to view the same footage three times to rate a different type of engagement each time: emotional, physical (Jones et al., 2015) and musical (Camic, Williams & Meeten, 2011). This was to help explore different aspects of engagement and identify an operational definition of it for the VASE.



*video available on YouTube

Figure 1. *Preliminary version of VASE*

Development stage

Face validity. A preliminary version of the VASE was developed by the Created out of mind team and uploaded onto an android tablet (Figure 1). This was field-tested with six healthy adult volunteers from the general public. The volunteers were asked to watch one of two YouTube videos preloaded onto the VASE app. Both video clips were around three minutes in duration and they consisted of a musician delivering a music intervention to a person with advanced dementia. Feedback was sought using open-ended questions on the usability of the app, and whether explanations about what behavioural expressions were regarded as engagement were adequate, and brief descriptions of their decision-making process during rating.

Verbal feedback were incorporated into the earlier version of the tool so that the VASE could be refined and revised. A thematic analysis was used to identify common patterns and categories about what engagement of PWD looks like (Appendix B). Braun and Clarke's (2006) six-stage approach was adopted when completing the analysis.

Content validity. Following open-ended feedback from the six volunteers, thematic analysis results were reviewed by the research team of three interdisciplinary experts (musician practicing music within residential care (JW), neuropsychologist (SC), clinical health psychologist (PC)) and a trainee psychologist (DL), in order to revise and create a protocol that would more accurately reflect the rating of engagement in the VASE. After the protocol was completed and adjustments made to the tool the final version was further tested by two volunteers.

Experimental stage

The experimental stage was conducted in two parts. The first part involved recording participatory music sessions (MfL and PL) involving PWD with severe dementia living in a care home. The second part involved recruiting raters to rate brief video excerpts from the participatory music sessions using the VASE. Each part of the study, the setting, participants and procedures, are described below.

Part 1.

Setting. Video recording of MfL and PL sessions were made in a London care home.

Participants. Eight PWD participated. Recruitment criteria: (i) a confirmed diagnosis of dementia at; (ii) an advanced stage (clinical dementia rating of 2–3) rated by staff; (iii) aged 60 or above; and (iv) able to sit in a room for an hour. PWD that had (i) a clinical dementia rating of below 2; (ii) severe hearing difficulties that cannot be corrected, even with a hearing aid; or (iii) disruptive behaviour during group activities in the care facility (e.g. aggressive behaviour) were excluded. These criteria were screened by care staff at the care home and verified by the researcher.

The intervention (MfL) and passive listening (PL) conditions. The MfL programme is one type of music-based intervention used for advanced dementia. It is an interactive music programme that was designed to promote better quality of life for PWD in residential care (Music for Life, 2014); it takes place for one hour a week over a course of eight weeks. Each week specially trained musicians facilitate and attempt to establish and enhance communication with the PWD through improvisational music and activities. The passive music listening (PL) session was held once prior to the beginning of the intervention. To maintain the experimental conditions, the controls listened to pre-recorded music that was similar to that used in the intervention sessions and played by the same musician. The

settings, length of session and number of musicians and care staff present at both intervention and control conditions were equivalent.

Procedures. The experimental stage first involved a one-hour control session consisting of listening to recorded music similar in style to the upcoming intervention sessions and played by the same musicians. This was followed one week later at the same time and in the same location by the start of the eight-week MfL intervention. At the beginning of each session, a 360-degree Fly video camera ® was placed in the middle of the room; this camera, which is smaller than a tennis ball, uniquely captures continuous 360-degree recording, making it ideal for use in groups. Videos were edited by an independent video editor into 30-second segments and 12 from the control and 4–5 from each of the intervention sessions were chosen using a table of random numbers. In each segment one of the eight participants were randomly chosen for the raters to specifically focus on and this was indicated with a yellow arrow (Figure 2). The 48 segments from the control and intervention sessions and PL session were then edited into a single 25-minute-long video. The order of the clips in the video was also random. Two more videos were made in the same manner as the first one, where the order of the clips was again randomly assigned to remove potential order effect. Videos were then transferred onto password protected and encrypted tablet computers that had preinstalled the VASE via a usb cable**.**

Part 2.

Setting. The video clips were then viewed using the tablets by the raters and scored in secure and non-public places. This included university and research organisations meeting rooms, a public library private meeting room and a charity's office.

Participants. Opportunity sampling was used to recruit professional and non-professional raters in Hong Kong and the UK through emails and face-to-face contact. Six professionals were included as raters (clinical and neuro-psychologists, nurses and dementia

charity managers). The inclusion criteria were as follows: (i) work in a health-related discipline; (ii) aged 18 or over; and (iii) one or more years of experience working with PWD. Thirty-one people from the general public (non-professionals) aged 18 and above, who had not worked clinically with PWD, were recruited.

Sample size. In accordance with the Medical Research Council (MRC) framework (2000), this was a feasibility study. Lancaster, Dodd and Williamson (2004) recommended an overall sample size of 30 for feasibility studies. Table 2 presents the demographic characteristics of the raters.

Table 2. *Demographic characteristics of raters*

	Total (N = 37)	
	N	Percent
Age (mean±SD)	38.2±2.69	
Gender		
Male	12	32.43
Female	25	67.57
Education		
High school or lower	2	5.40
Undergraduate	13	35.13
Master's	17	45.95
PhD or higher	5	13.51
Ethnicity		
White British	15	40.54
White other	4	10.81
Asian	17	45.95
Other	1	2.70
Participation in singing group		
Yes	23	62.16
No	14	37.83
Currently in singing group		
Yes	9	24.32

No	28	75.68
Experience with playing musical instrument		
Yes	17	45.95
No	20	54.05
Experience with PWD		
Yes	22	59.46
No	15	40.54
Type of Raters		
Professional	6	16.22
Non-professional	31	38.78

SD = standard deviation

Procedures. Thirty-seven participants, including professionals (n = 6) and people from the public (non-professional) (n = 31), were recruited as raters for this stage. Ratings were undertaken in secure and private locations (e.g. meeting rooms of universities, libraries and a research organisation). Once consent forms were signed, the raters were requested to fill in demographic information and were given the protocol on the categories of engagement to read. They were then given a password-protected tablet and over-ear headphones to complete a series of rating scales while watching the video recording of the MfL group. Research personnel were present to supervise this process and answer questions. Inter-rater reliability was tested to establish the consistency of the final version of the VASE. The researcher was present during all video viewings.

Data analysis

At this stage, raw data from the VASE, including the time and ratings were entered into EXCEL. The scores were then rounded to the nearest second. Data was then transferred into SPSS version 23 for data analysis.

Reliability. An intra-class correlation coefficient (ICC) was used to determine the consistency of coefficients across all raters. This was also used to assess the engagement

scores across the three videos, the conditions (PL and MfL session) and the raters (profession and non-professionals). Spearman's correlation was used to evaluate the inter-correlations between the professionals, as well as between the non-professionals.

Mixed model analysis. Multilevel linear modelling (MLM) was used to investigate whether extraneous variables, including age, gender, experience working with PWD and experience with music, might have an effect on the rater's engagement rating. Since one participant would appear in multiple clips and they were rated by each rater at these different time points, MLM takes account of the dependencies by estimating variance associated with group (e.g. raters) and differences in average response (intercepts). The model adopted for data analysis in this study regards intercepts and/or slopes to be random effects. All of the analyses were conducted using SPSS Version 23 and the level of the significance was set to 0.05.

Ethical considerations

Ethics approval was granted by a Canterbury Christ Church ethics panel and also approved by the charity where the MfL group was conducted (Appendix C).

***People with dementia*.** This study followed the guidelines for working with people who are unable to directly provide informed consent. An information sheet (Appendix D) was given to caregivers and consent obtained from the participants' primary caregiver (Appendix E), who has the legal authority to give consent. Before each video recording, centre staff and musicians would also verbally remind the PWDs that the sessions were being recorded for the purpose of this research and offer them the opportunity to withdraw from the recording.

Raters. As the raters would be watching a recording of a vulnerable population the information sheet (Appendix F) and consent form (Appendix G) specifically highlighted that the video would only be used for the purpose of research and possible identifying details

should not be shared. The consent form also reminded the raters that they should not disclose the name or identifying information of the PWDs viewed in the video.

Video recordings. These were transferred from the camera directly to encrypted files in a password-protected computer tablet and the video data was erased from the camera.

Results

Development stage

Face validity

Based on the feedback from the volunteers watching the YouTube videos, some adjustments were made; this included adjusting the size of the videos to 850mm x 500mm and adjusting the font size. The volunteers reported finding it difficult to distinguish the different types of engagement. Some volunteers also expressed that watching the same video three times made them lose interest and, as a result, they found it difficult to concentrate during the repeat viewings. Consequently, the primary researcher and Created Out of Mind team held a meeting to discuss the pilot feedback and findings. The team acknowledged that perhaps asking raters to rate the same video three times in relation to different types of engagement led to fatigue and for non-professionals, distinguishing types of engagement was difficult. Rating procedures being perceived as repetitive or confusing would defeat the purpose of creating a user-friendly engagement tool. The purpose of the tool was to examine PWD's ability to engage in a group rather than to find out what type of stimulus caused engagement or to differentiate types of engagement. Therefore, it was decided that the VASE should be used to rate an overall state of engagement. After making the revisions on the app, a further field-test was carried out with five more volunteers to ensure that the final version of the VASE was suitable (Figure 2).

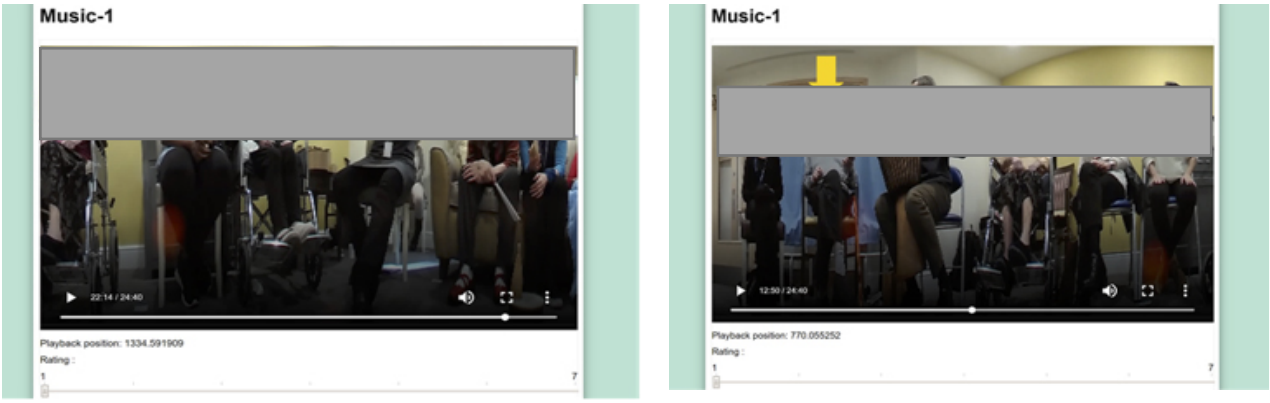


Figure 2. Snapshots of the *Final version of the VASE*

In addition, based on the questions related to aspects that made raters consider a PWD to be engaged in the group, a thematic analysis was carried out using the interview data gathered from volunteers. The interviews were transcribed and four main categories (patterns) and 12 behavioural expressions of the categories were identified (See Table 3).

Table 3. *Categories and behaviour expressions of engagement*

Categories	Behavioural expressions
Facial expressions	i. Mouth and lip movement
	ii. Eyebrows movement (e.g. closed their eyes or raise their eyebrow)
	iii. Facial changes (e.g. Neutral look, smile)
Bodily movement and verbal articulations	i. Large and subtle bodily movement (e.g. Hands and feet taping, nodding, clapping, moving with music)
	ii. Verbally responding (e.g. singing, talking, mouth mumbling)
	iii. Interacting with instruments (touching the instrument, playing with the instruments, making music)
Attention and awareness of activity	i. Attention to Stimulus (musician, other participants) undistracted eye contact
	ii. Playing an instrument
	iii. Moving along with music
Emotional response	i. Pleasure and enjoyment as indicated by smiles and a look of contentment
	ii. At ease look (looks as if s/he was relaxed)
	iii. Sad or anxious look (appear agitated, e.g., eyes down casted like in moment of unhappiness; tapping his/her fingers as in people who are anxious)

Content validity

Four main categories were derived for the initial protocol based on the qualitative data analysis. The four categories included: a) facial expressions; b) bodily movement and verbal articulations; c) attention and awareness of activity; and d) emotional responses. The initial protocol was much briefer, offering little description of the categories. There was 100% agreement from the experts, indicating that the categories and their corresponding behavioural expressions are a good representation of engagement in PWD. However, the experts also highlighted the importance of offering some examples of behaviours relating to each category. They specifically considered that explanations should include behavioural expressions that were not easily picked up as a sign of engagement. For instance, one of the experts spoke about PWD experiencing what other people might describe as “negative” emotions, such as sadness, where they might be considered to be emotionally “moved” by the

music. The experts further commented that there is a need to highlight that sometimes eyes being closed, or even a natural look, can be a sign of a person being engaged. In addition, the professionals rating would also be dependent upon different cues, such as the context of the situation, the rater's own experience, and the rater's understanding of the group. Consequently, some experts proposed that the protocol should not be a rigid manual; rather, it should simply provide a reference for what engagement is and allow a certain amount of ambiguity and openness towards a rater's own interpretation. The appropriateness of the rating would be determined in the third stage of inter-rater reliability testing. To see if there was consistency and agreement, different raters' scoring of the same individual at the same time were statistically tested to examine if variances existed. Consequently, a statement about there being no right or wrong answer was added to the brief description. This process resulted in the final version of the VASE protocol (Appendix H).

Experimental stage

Observing raters during the viewing sessions, and from informal comments made by the raters, it was apparent that they had a variable delay in making their rating as each new video appeared and they appraised the scenario. As a result, the first five seconds of each clip were excluded from the analysis of the rating data gathered.

In total, each rater watched 1,200 seconds of clips, of which 300 seconds were PL and 900 seconds were MfL sessions. The VASE was tested across 37 raters, with three sets of videos that consisted of the same 48 clips, but each edited into a different order (M1, M2 and M3). The distribution of rating and average rating over time was analysed based on raters' characteristics, which are listed in Table A1 in Appendix I. The analysis included the rating provided by raters overall, video conditions (MfL versus PL), gender, videos watched (different clip order: M1, M2 and M3), professional or non-professional, with or without

experience in looking after PWD, and experience of musical instrument and singing group. The average rating of each MfL session was also reported (Appendix I).

The mean rating score of each individual rater under MfL and PL conditions are reported in Table 4. A non-parametric Man-U-Whitney test was used to analyse whether there were any differences in rating between the two conditions by each rater. The results showed that most of the raters (36 out of 37 raters) gave significantly lower ratings for PL than MfL sessions. Higher ratings were observed for MfL than PL sessions irrespective of the order in which the videos were viewed and rated (M1, M2 or M3) or their professional/non-professional status (see Table 4). It should be noted that the mean MfL ratings shown in Tables 4 and 5 are mean scores of MfL sessions 1 to 8 (see Figure 2). Differences between MfL and PL mean ratings suggest that the inter-class correlation should be analysed separately for MfL and PL sessions.

Table 4. Scoring by *Individual rater as analysed by MfL and PL conditions*

Raters	Video	Total				PL				MfL				Mann-Whitney U test
		N	Mean	SD	Median	N	Mean	SD	Median	N	Mean	SD	Median	
1	M1	1200	2.4	1.41	2.0	300	1.7	0.78	2.0	900	2.7	1.49	2.0	10.752 ***
2	M1	1200	4.1	2.18	4.0	300	4.5	2.38	5.0	900	3.9	2.09	4.0	4.297 ***
3	M1	1200	3.9	2.03	4.0	300	2.7	1.91	2.0	900	4.3	1.92	4.0	11.211 ***
4	M1	1200	3.2	1.56	3.0	300	2.4	1.50	2.0	900	3.5	1.48	3.0	10.605 ***
5	M1	1200	3.8	2.24	3.0	300	2.2	1.55	2.0	900	4.3	2.19	4.0	14.089 ***
6	M1	1200	4.1	2.33	4.0	300	2.8	2.14	2.0	900	4.5	2.23	4.0	11.298 ***
7	M1	1200	3.0	1.62	3.0	300	2.1	1.22	2.0	900	3.3	1.61	3.0	12.340 ***
8	M1	1200	1.8	1.23	1.0	300	1.2	0.44	1.0	900	2.1	1.32	1.0	11.539 ***
9	M1	1200	4.3	2.08	5.0	300	2.7	1.64	2.0	900	4.8	1.93	5.0	15.319 ***
10	M1	1200	4.5	1.99	5.0	300	3.0	1.33	3.0	900	5.1	1.91	6.0	15.564 ***
11	M2	1200	3.7	1.91	4.0	300	2.4	1.64	2.0	900	4.2	1.78	4.0	14.244 ***
12	M2	1200	3.1	1.33	3.0	300	2.3	1.19	2.0	900	3.3	1.27	3.0	11.143 ***
13	M2	1200	1.9	1.40	1.0	300	1.3	0.71	1.0	900	2.1	1.51	1.0	8.628 ***
14	M2	1200	1.6	0.88	1.0	300	1.3	0.57	1.0	900	1.7	0.94	1.0	6.692 ***
15	M2	1200	4.5	1.97	5.0	300	3.2	1.88	3.0	900	4.9	1.80	5.0	12.806 ***
15	M2	1200	5.5	2.12	7.0	300	4.7	2.46	6.0	900	5.8	1.93	7.0	5.942 ***
17	M2	1200	2.8	1.78	2.0	300	1.8	1.08	1.0	900	3.1	1.85	3.0	11.570 ***
18	M2	1200	3.8	2.16	4.0	300	2.6	2.03	1.0	900	4.2	2.06	4.0	11.314 ***
19	M2	1200	5.1	1.68	6.0	300	4.0	1.86	5.0	900	5.5	1.43	6.0	12.870 ***
20	M2	1200	2.6	1.84	2.0	300	2.7	2.08	2.0	900	2.6	1.75	2.0	0.962
21	M2	1200	3.1	1.61	3.0	300	1.8	1.03	1.0	900	3.5	1.52	3.0	17.696 ***
22	M2	1200	2.7	2.16	1.0	300	1.7	1.70	1.0	900	3.0	2.19	2.0	11.201 ***
23	M3	1200	2.4	1.56	2.0	300	2.1	1.17	2.0	900	2.5	1.65	2.0	2.928 ***
24	M3	1200	3.4	1.88	4.0	300	2.3	1.49	2.0	900	3.8	1.85	4.0	12.061 ***
25	M3	1200	2.9	1.16	3.0	300	2.1	0.89	2.0	900	3.1	1.12	3.0	13.274 ***
26	M3	1200	2.4	1.60	2.0	300	1.7	0.85	1.0	900	2.7	1.72	2.0	8.709 ***
27	M3	1200	3.4	1.79	3.0	300	2.2	1.14	2.0	900	3.8	1.79	4.0	13.809 ***
28	M3	1200	3.2	2.07	2.0	300	1.8	1.08	1.0	900	3.7	2.09	4.0	14.341 ***
29	M3	1200	3.9	2.02	4.0	300	2.1	1.27	1.0	900	4.5	1.84	5.0	18.253 ***
30	M3	1200	4.4	1.77	5.0	300	3.1	1.72	3.0	900	4.9	1.53	5.0	14.695 ***
31	M3	1200	1.9	1.23	1.0	300	1.6	0.93	1.0	900	1.9	1.30	1.0	3.200 ***
Raters who were professionals.														
A	M1	1200	4.0	1.92	4.0	300	2.8	2.00	2.0	900	4.4	1.72	4.0	11.732 ***
B	M1	1200	4.2	1.74	5.0	300	2.8	1.57	3.0	900	4.7	1.52	5.0	16.059 ***
C	M2	1200	4.6	2.18	5.0	300	3.4	2.21	3.0	900	5.0	2.02	6.0	10.239 ***
D	M3	1200	3.6	1.68	4.0	300	2.4	1.40	2.0	900	4.1	1.55	4.0	14.502 ***
E	M3	1200	3.6	1.73	4.0	300	2.2	1.29	2.0	900	4.1	1.57	4.0	16.928 ***
F	M3	1200	3.0	1.56	3.0	300	1.9	0.84	2.0	900	3.3	1.58	3.0	14.404 ***

*** $p < 0.005$

SD=Standard deviation, MfL=Music for Life, PL=Passive music listening, M1, M2, M3=Video order

Table 5. *Influence of video rating order and professional status upon engagement ratings for MfL and PL conditions.*

Video (order)	PL				MfL				Mann-Whitney U test
	N	Mean	SD	Median	N	Mean	SD	Median	
M1	300	2.6	1.14	2.5	900	4.0	1.29	3.8	14.636 ***
M2	300	2.5	1.05	2.4	900	3.8	1.21	3.8	14.243 ***
M3	300	2.1	0.83	2.1	900	3.5	1.17	3.6	17.382 ***
Professionals	300	2.6	1.25	2.5	900	4.3	1.15	4.2	17.443 ***
Non-professionals	300	2.4	0.94	2.2	900	3.7	1.10	3.7	15.413 ***

*** $p < 0.005$

SD=Standard deviation, MfL=Music for Life, PL=Passive music listening, M1, M2, M3=Video order

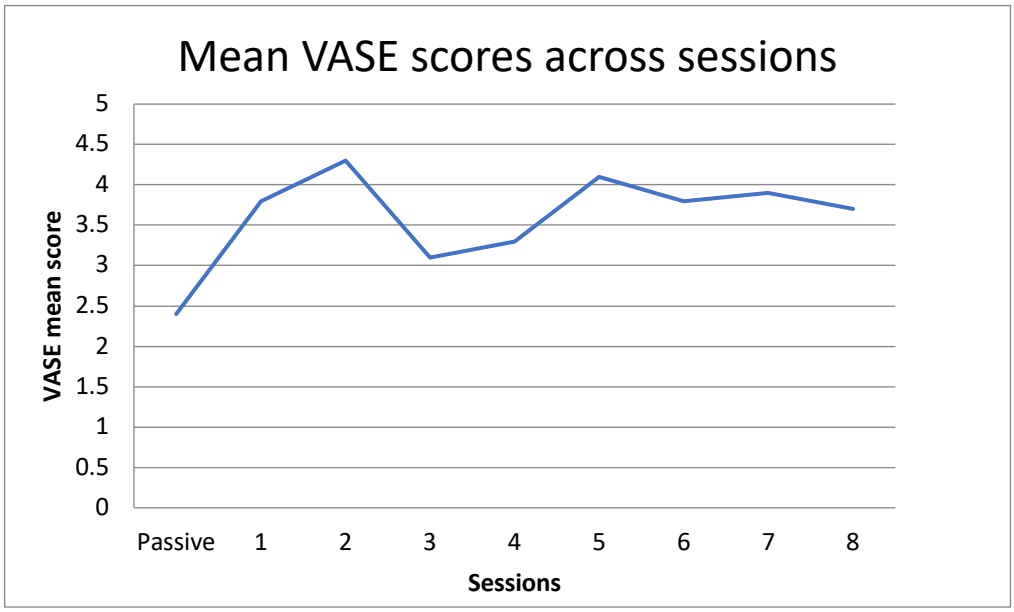


Figure 2. Mean *VASE scores across sessions*

Reliability.

Inter-rater reliability. In our study, the Intra-class Correlation Coefficient (ICC) was used to measure the consistency of raters in the rating of the engagement of the target participants over 48 video clips. According to Koo and Li's criteria (2016), values less than 0.5 are considered to have poor reliability, 0.5 to 0.75 are considered to have moderate reliability, 0.75 to 0.9 is considered to have good reliability and ICCs greater than 0.9 are considered to have excellent reliability.

The ratings of the three videos (different clip order) of M1, M2 and M3 and the ratings for MfL and PL sessions were assessed for their inter-rater reliability. The values of ICC in the MfL sessions ranged between 0.841 and 0.876, while the corresponding values in the PL session ranged between 0.812 and 0.883. All of the values were greater than 0.8, which is considered to reflect good reliability. Comparing the ratings of professional experts and the general public, the values of ICC in both groups indicated an excellent level of reliability (MfL: 0.881; PL: 0.938). These results help accept hypothesis 1 and indicate that

the professional experts and the general public showed strong agreement in both conditions (MfL and PL). The mean ICC values suggest the reliability levels were good among the professionals (MfL: 0.854; PL: 0.775) and excellent among the non-professionals (MfL: 0.918; PL: 0.916), although the different group sizes (professional: N = 6; non-professional: N = 31) should be noted in interpreting these values. The findings accept hypothesis 2 and suggest that there is a correlation between the VASE rating among the dementia experts, as well as among non-professionals.

The mean rating scores of the raters across videos (M1, M2 and M3) were assessed (Table 6). For M1 and M2, the values of ICC were greater than 0.9 for both M1 and M2, indicating excellent reliability. The values of ICC for M1 and M3, as well as M2 and M3, showed good reliability with a few of them achieving moderate reliability.

Table 6. *Intraclass correlation coefficient (ICC)**

	Value	95% CI
Inter-rater reliability		
M1	0.896	(0.869 , 0.916)
MfL	0.876	(0.839 , 0.903)
PL	0.867	(0.823 , 0.898)
M2	0.861	(0.804 , 0.897)
MfL	0.841	(0.770 , 0.885)
PL	0.812	(0.747 , 0.858)
M3	0.901	(0.876 , 0.920)
MfL	0.869	(0.828 , 0.898)
PL	0.883	(0.858 , 0.905)
Professional vs Non-professional	0.920	(0.637 , 0.967)
MfL	0.881	(0.258 , 0.957)
PL	0.938	(0.909 , 0.956)
Professional	0.850	(0.811 , 0.878)
MfL	0.854	(0.812 , 0.886)
PL	0.775	(0.717 , 0.818)
Non-professional	0.934	(0.920 , 0.944)
MfL	0.918	(0.898 , 0.935)
PL	0.916	(0.897 , 0.931)
Inter-retest reliability (across groups*)*		
M1 vs M2	0.951	(0.939 , 0.960)
MfL	0.936	(0.913 , 0.951)
PL	0.959	(0.948 , 0.967)
M1 vs M3	0.804	(0.713 , 0.859)
MfL	0.710	(0.617 , 0.774)
PL	0.867	(0.498 , 0.943)
M2 vs M3	0.786	(0.743 , 0.820)
MfL	0.679	(0.630 , 0.722)
PL	0.866	(0.523 , 0.941)

* All of ICC values are significant with $p < 0.001$

CI=Confidence interval, MfL=Music for Life, PL=Passive music listening, M1, M2, M3=Video order

Spearman's correlation coefficients between the professionals were evaluated. The results show that the coefficients ranged from 0.353 to 0.72 (Table 7). In our study, there was a high correlation between experts E, F and G, which indicated that they had a high level of agreement in their ratings. The correlation coefficients among experts B, E and F were greater than 0.5, which indicates that they had a moderate correlation. The correlation between experts D, E, F and G, as well as that between experts A, B and C was also moderate. On the other hand, experts A and C had a weak correlation with D, E, F and G. All the correlation coefficients indicated a statistically significant degree of correlation. Table A2 in Appendix J shows the correlation coefficients between the non-professionals. The results show that 89.5% (416 out of 465) of the correlation coefficients between raters are significantly correlated at 5% level. In general, the findings support hypothesis 2, indicating that there is a correlation between the VASE rating among professionals, as well as non-professionals.

Table 7. *Spearman's correlation coefficient in rating for professional*

Raters	A	B	C	D	E	F
A	1					
B	0.641***	1				
C	0.535***	0.644**	1			
D	0.371***	0.403***	0.402***	1		
E	0.441***	0.599***	0.466***	0.575***	1	
F	0.353**	0.571***	0.396**	0.500***	0.720***	1

*** $p < 0.005$

Mixed model analysis

Data were averaged into five five-second stages and recoded into new variables called "stages". This variable categorised each five seconds of a clip into stages: stage 1 (6s–10s), stage 2 (11s–15s), stage 3 (16s–20s), stage 4 (21s–25s) and stage 5 (26–30s) (see Figure 3).

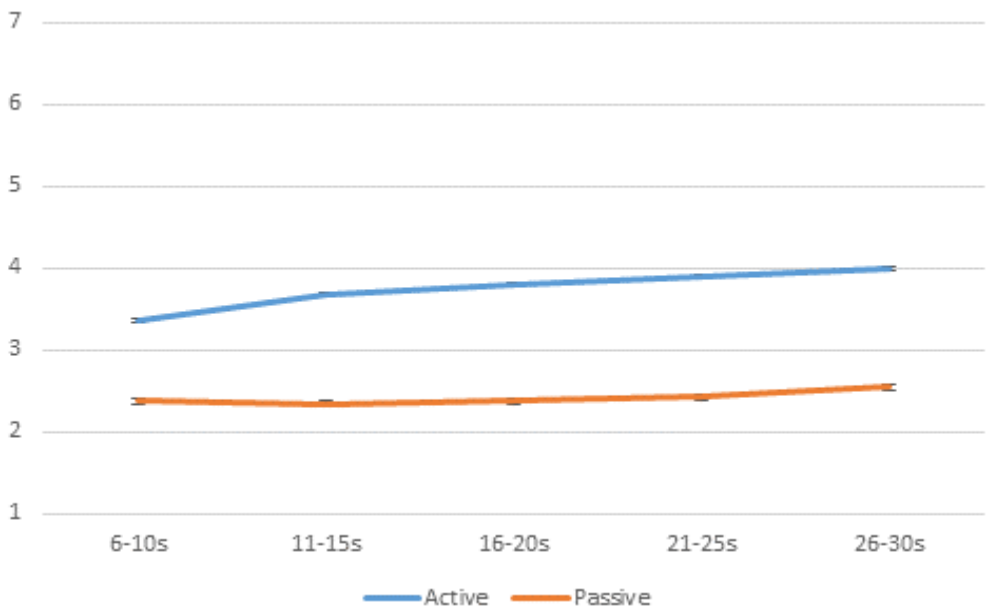


Figure 3. Engagement ratings between MfL and PL sessions across stages

MLM was used to investigate the relationship between the rated level of engagement and a number of variables, including condition type (active [MfL]/passive [PL]), session number (1–8), within-video "stage" (1–5), and rater characteristics, such as age, gender, professional status (professional/non-professional), experience of playing an instrument, experience of singing in a choir, presence in one of the MfL sessions, and experience of PWD (all were dichotomous 'yes/no' ratings).

In our hypothetical model, MfL or PL listening, sessions, engagement across the five "stages", raters' profile such as age, gender, profession, playing an instrument and experience of singing in a choir, experience in a choir or not, and experience with PWD were entered as fixed effect (without interaction term), while rater and PWD recorded in the videos were

entered as random effect, which is based on the hypothesis that there would be a difference in the relationship between the level of engagement and raters, as well as the PWDs.

A full model (-2 log likelihood = 167030.821) that includes all of the variables is significantly better than one in which only the intercepts are included (-2 log likelihood = 173760.602), with λ^2 (10, N = 44400) = 6729.781, $p < 0.001$. Thus, inclusion of all variables improved the model beyond that produced by considering variability in raters and participants. This significantly lower level of chi-square in the full model provided the rationale that MLM should be used. Among the 10 predictors selected into the model, half of them were significantly associated with the level of engagement (Table 8). These five variables were stage ($\beta = 0.12$, $p < 0.005$), MfL or PL condition ($\beta = -1.43$, $p < 0.005$), session ($\beta = 0.02$, $p < 0.005$), age ($\beta = -0.03$, $p < 0.005$) and profession ($\beta = 0.97$, $p < 0.05$). Results showed that ratings differed significantly across the five "stages", and on average went up by 0.12 per stage indicating that raters' ratings were changing over time, which might suggest PWD were more responsive when session progressed. Engagement was rated significantly higher for the MfL sessions than the PL session, on average 1.43 points on the rating scale. In addition, there were significant differences in engagement ratings across the sessions. Ratings were recorded as higher at the latest session than the earlier session, with on average 0.02 points difference per session, which might not be clinically significant. Age of rater was also a significant factor, with younger raters providing higher ratings than their older counterparts. Professional raters provided significantly higher ratings. Playing an instrument and having experience of singing in a choir did not seem to influence ratings. Similarly, ratings did not significantly differ if raters had experience of working with PWD.

Table 8. *Estimates of fixed effects*[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	4.844	0.501	37.203	9.672	0.000	3.830	5.859
stage	0.122	0.005	44101.510	23.105	0.000	0.111	0.132
Condition	-1.435	0.026	44353.786	-54.713	0.000	-1.486	-1.383
Session	0.024	0.004	44201.130	5.350	0.000	0.015	0.033
Age	-0.031	0.009	37.000	-3.319	0.002	-0.050	-0.012
Profession	0.970	0.382	37.000	2.541	0.015	0.196	1.743
Rater' Gender	-0.073	0.264	37.000	-0.276	0.784	-0.608	0.462
Instrument	-0.349	0.360	37.000	-0.970	0.338	-1.079	0.380
Presently in Sing group	0.746	0.372	37.000	2.006	0.052	-0.007	1.500
Sing group	-0.470	0.309	37.000	-1.519	0.137	-1.097	0.157
PWDexp	-0.614	0.384	37.000	-1.600	0.118	-1.391	0.163

a. Dependent Variable: Rating.

Table 9 shows the random effects of the model. It was found that there was significant variability in the ratings given by different raters ($p < 0.001$), as well as significant variability in the rating of the PWD between raters ($p < 0.001$). There was also significant residual variance after taking into account all effects in the model. This residual variance might indicate that the model requires more variables. The residuals of the model were tested with a Q-Q plot: it was found that the residuals followed a normal distribution and it was thus concluded that the normality assumption of the model is supported (Appendix K).

Table 9. *Estimates of covariance parameters*[a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		2.456	0.017	148.495	0.000	2.424	2.489
Intercept [subject = Rater * VidSubject]	Variance	0.679	0.062	10.939	0.000	0.567	0.812
Intercept [subject = Rater]	Variance	0.424	0.119	3.556	0.000	0.245	0.736

a. Dependent Variable: Rating.

Discussion

The aim of this feasibility and validation study was to develop a publicly available user-friendly engagement tool for use in the assessment of people with advanced dementia. The study helped better understand engagement in PWD by creating an operational definition of the concept. It also showed the possibility of adopting VASE as a measure for the aforementioned use particularly in-group settings and shed the light of the possibility of using a dynamic, continuous ratings scale that captures concepts that entails dynamic changes moment-by-moment, like that of engagement. In the study, face validity was obtained based on the opinions of volunteers from the general public. Thematic analysis of interview data was used to construct an operational definition of engagement. Inter-rater-reliability was documented, and strong agreement was found in some conditions.

Hypothesis 1 testing the correlation between VASE rating by dementia professionals and non-professional people is accepted. ICC indicated that VASE has good to excellent agreement between the two samples of professionals and general public. Yet, the MLM suggested that overall, professionals generally rated a significantly higher score than the general public (average of 0.970). This could be a result of professionals having more clinical knowledge and understanding about engagement of people with advanced stage dementia than the general public. Without as much experience with advanced dementia the general public might not be as aware of the challenges PWD experience in their capacity to verbally and physically respond; therefore, they may have rated lower on the VASE scale.

Hypothesis 2 tested inter-rater reliabilities and is accepted. When looking at inter-rater reliability across professionals high to moderate ICC were found in general with some inconsistent results between raters. This inconsistency might be the result of experts working in different fields and different settings. For example, those who had higher agreement with

each other were psychologists and clinicians working in community settings and have experience running groups for PWD. The others work in more acute hospital settings; consequently, they have less experience of residential group interventions. As a result, the professionals working in acute wards might potentially have different understandings of what engagement looks like for PWD than those working in residential care. Indeed, a person's belief system, worldview and reality are often constructed based on their experience (Koltoko-Rivera, 2004). This was difficult to account for and capture in the VASE and it raises questions about whether the current VASE is overly reductionist in capturing such a complex concept as engagement. On the other hand, high agreement was found from non-professionals, further supporting the reliability of VASE among non-professional raters.

Hypothesis 3 tested whether VASE will be able to differentiate the level of engagement between MfL and PL, and is accepted. The ICC and non-parametric testing results suggested that the VASE has good to excellent reliability in differentiating MfL from PL sessions. These results echo the findings from MLM where raters generally rated the control condition an average of 1.435 points lower. This result is perhaps unsurprising as active MfL music activities are more dynamic and interactive than passive music listing, where participants in the former tend to react and respond to each other and the musician(s) also making the music using different musical instruments (American Music Therapy Association, 2015).

Hypothesis 4 assumed that the extraneous variables would not affect rating. Hypothesis 4 is rejected. The MLM findings suggest that ratings differ significantly across stages (five seconds). This suggests that the level of engagement rated in the video changes over time and raters are using the VASE in a dynamic fashion. Another promising result was that the order in which raters rated the segments (M1 versus M2 versus M3) had no effect on

rating value. This is consistent with the between-group ICC result and suggests that the VASE rating scale is not affected by order effect.

The rater's age seemed to have an effect on the rating score. During data collection some participants aged in their 70s and above expressed that they found it difficult to use the tablets and a 25 minute video was too long for them. Cornish and Dukette (2009) stated that the average maximum attention time for adults is around 20 minutes, therefore, future research should consider the optimal length of time to use the VASE. Apart from familiarity with technology and attention span, other factors such as decline in processing speed (Eckert, Keren, Roberts, Calhoun & Harris, 2010) and response selection time (Woods, Wyma, Yund, Herron & Reed, 2015), have been found to be associated with ageing. To overcome this issue it is worth considering using tests for reaction time commonly used in computerised neurocognitive tests to learn about raters' baseline reaction times (Donald et al., 2015), or revise the program in such a way that permits a longer processing duration or to pause and rewind.

Strength and weaknesses

The present study developed a new rating tool and examined its validity. This study explored the possibility of using "non-symptomatic" concepts such as engagement to understand PWD's response in interventions. Engagement with others and involvement in activities are important for various dimensions of health and wellbeing for PWD (Benveniste, Jouvelot, Pin & Pequignot, 2012; Moyle et al., 2013; Robinson, MacDonald, Kerse & Broadbent, 2013). A validated tool for the assessment of engagement will be useful to researchers and clinicians to better understand the effect of interventions for PWD and those who might have difficulties verbally expressing themselves. The examination of the processes during an intervention is crucial in helping professional and family carers to learn

about the participants' responses, and to gauge clinical benefits. VASE is therefore a useful tool in that sense. When such a tool is user-friendly and publicly available, it will have wider applicability to all those who are involved in dementia care.

Secondly, the study made a unique contribution as it was the first known study to capture moment-by-moment changes of engagement that take place during the intervention, enabling raters to continuously make ratings as they observe changes in the video. The VASE is also non-intrusive and it does not require raters to be present due to the use of a previously recorded video. Videos can be reviewed and re-rated again by the same viewer or different viewers, enabling multiple raters to cross-track their engagement scores. Most importantly, the VASE can record the exact time that changes in engagement occur during a group intervention. This allows raters to know which particular activities stimulate different levels of engagement for particular individuals. This could potentially enable clinicians or carers to tailor specific activities for PWD in order to promote person-centred care. This tool might allow non-professional caregivers who are living away from the PWD to be involved in tailoring activities for their family member.

One of the limitations of the tool is that this was a feasibility study and the sample size of raters was relatively small. Furthermore, as sampling was opportunistic, there is the possibility of sampling bias (Stasser & Titus, 1985), and people who participated in this experiment might have different attitudes and understandings about engagement from those who did not participate in the study. The recruitment criteria also meant that people who have a good understanding of dementia (e.g. academics and dementia carers), but are not health professionals, were considered general public, possibly confounding the study results.

Criterion and construct validity were also not established. It was not possible to establish criterion validity as the VASE is a single item continuous rating scale. However further research could determine construct validity to enable better confidence that the tool is

operating theoretically as expected (Gaugler, Hobday & Savik, 2013). Test-retest reliability was also not established due to time constraints. Further, rating participants on one single music group does not ensure the results are generalisable to other settings.

Lastly, during the discussion on the use of the tool, some experts opined that in order to be able to fully understand engagement in a session, it is necessary to be present in the group and be "immersed in that atmosphere". Some experts also commented that ratings of engagement might be subjective and based on the raters' understanding of, and familiarity with, the subject they are rating. As a result, there might be factors that this study had not considered. As this is the first cycle in its development, further work is needed before we can be confident in its reliability.

Research Implications

Future research could review the current version of VASE and investigate possible adjustments of the tool, such as trialling with ratings based on different types of engagement and reviewing the duration of the rated video segments. Further research should also investigate the difference between the general public and professionals in their understanding of engagement of PWD living in care homes. The protocol could then be revised to consider these differences. With a larger sample size different validity and reliability tests could be used, such as criterion and construct validity. Test-retest reliability could also be considered with the same rater re-rating the video.

The VASE adopted a seven-point rating; seven-point ratings have been previously recommended as a good multi-point scale in preference to a five-point scale (Lewis, 1993). However, the results of the mixed model suggest that the mean difference (standard estimate) between conditions, such as session and conditions (control and intervention) was small,

further examination on the sensitivity and specificity of the tool is needed to better understand the statistical and clinical significance of outcomes.

In the MLM, residual (unexplained) variance reached a significant level. This suggests there are currently other variables that have not been considered and more variables need to be incorporated in order to build a better model (Heck & Thomas, 2000). Variables such as mood, cognitive ability, attitude towards the activity could be added. For the raters their awareness of dementia and attitudes towards dementia (Handley, Bunn & Goodman, 2017), and their age and baseline reaction time could also be investigated.

Further research could examine the use of VASE in other interventions regarded as beneficial for PWD such as cognitive stimulation (Orrell et al., 2017), art therapy (Deshmukh, Holmes & Cardno, 2018) and other types of music interventions.

Clinical Implications

The mixed findings suggest that further refinement of the VASE is needed before it can be used in clinical practice. Engagement as an outcome was not previously considered a worthwhile construct to measure in a dominant medical model of dementia care but it is now deemed valuable with person-centred (Kidwood, 1997; Sung & Chang, 2005) and relational (Greenwood, Loewenthal & Rose, 2002) approaches. Unlike symptom-based tools that measure the success of the intervention the VASE offers an alternative way of understanding PWD, investigating choice and interaction.

With revision, the VASE could be adopted in residential care settings by clinical psychologists and others to help understand levels of engagement with various different activities. This could enable care staff to assess whether particular activities are suitable for individuals. As the tool provides a time stamp in terms of noted changes in engagement, it could also assist group facilitators to identify particular activities or stimuli that support

higher levels of engagement and those that support less. This could enable facilitators to adjust their intervention based on the group or the individual's preference. Facilitators could also take the tool and results to other experts, and even clinical supervision, in order to make changes that would benefit PWD.

As a video analysis tool that is simple and user-friendly, the VASE could potentially be beneficial for inviting a wider support network to engage in the care of the individual. For example, the tool could enable family carers to view and evaluate activities that PWD participate in a care home without needing to be present in the group, allowing families the option of becoming more involved in PWD's care.

Lastly, if the measure is found to be valid, it could potentially be used for staff training, where examples of engagement and non-engagement can be identified and shared. Most importantly, the new observational tool could also enable us to gain a better understanding of the particular nuances and components of what makes an activity useful for this population, and potentially this could be applied to evaluations of other types of interventions and activities (e.g. museum object handling, approaches to self-care, family interactions) for PWD. Dynamic rating scales beyond engagement could also be adopted, applying them to other concepts in dementia care where observation of dynamic changes in PWD is required.

Conclusion

The feasibility and validation study results indicate that the current version of the VASE has good reliability in some areas. It still needs further investigation and adjustments for it to be a valid and reliable tool in measuring engagement of PWD in a group setting. Balancing the wish to develop a user-friendly tool and to capture a complex abstract concept such as engagement is challenging. It is encouraging that there is some evidence suggesting that the VASE is able to distinguish between the level of engagement of participants in two different

types of music activities (passive and active). With further adjustments and investigation, the VASE could be a useful tool in advancing dementia care. Improvements in assessments of the processes during an intervention will facilitate a better capturing of the concept of engagement, and would eventually benefit carers in the promotion of wellbeing of those who suffered from dementia.

References

Abraha, I., Rimland, J. M., Trotta, F. M., Dell'Aquila, G., Cruz-Jentoft, A., Petrovic, M., ... & Cherubini, A. (2017). Systematic review of systematic reviews of non-pharmacological interventions to treat behavioural disturbances in older patients with dementia. The SENATOR-OnTop series. *BMJ open*, *7*, e012759. doi: 10.1136/bmjopen-2016-012759

Algar, K., Woods, R. T., & Windle, G. (2016). Measuring the quality of life and well-being of people with dementia: A review of observational measures. *Dementia*, *15*, 832-857. dol: 10.1177/1471301214540163

Alzheimer's Society (2015, July). *Dementia 2015: Aiming higher to transform lives* Retrieved from https://www.alzheimers.org.uk/sites/default/files/migrate/downloads/dementia_2015_aiming_higher_to_transform_lives.pdf

Alzheimer's Society. (2018). *The dementia guide: Living well after diagnosis*. London: Alzheimer's Society.

Asan, O., & Montague, E. (2014). Using video-based observation research methods in primary care health encounters to evaluate complex interactions. *Informatics in Primary Care*, *21*, 161. doi: 10.14236/jhi.v21i4.72

Baird, A., & Samson, S. (2015). Music and dementia. *Progress in Brain Research, 217,* 207-235. doi: 10.1016/bs.pbr.2014.11.028

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3,* 77-101. dol: 10.1191/1478088706qp063oa

Braun, M. (2019). Management of Behavioral and Psychological Symptoms in Dementia. In Ravdom, D. L., Katzen. L. H., (Eds.) *Handbook on the neuropsychology of aging and dementia* (pp. 355-364). London: Springer.

Beard, R. L., Knauss, J., & Moyer, D. (2009). Managing disability and enjoying life: How we reframe dementia through personal narratives. *Journal of Aging Studies*, *23*, 227-235. doi: 10.1016/j.jaging.2008.01.002

Brooker, D., & Latham, I. (2015). *Person-centred dementia care: Making services better with the VIPS framework*. London: Jessica Kingsley.

Cabrera, E., Sutcliffe, C., Verbeek, H., Saks, K., Soto-Martin, M., Meyer, G., ... & Right Time Place Care Consortium. (2015). Non-pharmacological interventions as a best practice strategy in people with dementia living in nursing homes. A systematic review. *European Geriatric Medicine*, *6*, 134-15. doi: 10.1016/j.eurger.2014.06.003

Camic, P. M., Williams, C. M. & Meeten, F. (2011). Does a 'singing together group' improve the quality of life of people with a dementia and their carers? A pilot evaluation study. *Dementia, 12*, 152-171. doi: 10.1177/1471301211422761.

Carthey, J. (2003). The role of structured observational research in health care. *BMJ Quality & Safety, 12*, ii13-ii16. dol: 10.1136/qhc.12.suppl2.ii13

Cheston, R., & Ivanecka, A. (2017). Individual and group psychotherapy with people diagnosed with dementia: A systematic review of the literature. *International Journal of Geriatric Psychiatry*, *32*, 3-31. doi: 10.1002/gps.4529

Cornish, D & Dukette, D. (2009). *The essential 20: Twenty components of an excellent health care team*. Pittsburgh, PA.: RoseDog Books.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In Tashakkori, A & C. Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioural research* (pp. 209-240). Thousand Oaks, CA: Sage.

Cohen-Mansfield, J., Dakheel-Ali, M., & Marx, M. S. (2009). Engagement in persons with dementia: the concept and its measurement. *The American Journal of Geriatric Psychiatry*, *17*, 299-307. doi: 10.1097/JGP.0b013e31818f3a52

Cooke, M. L., Moyle, W., Shum, D. H., Harrison, S. D., & Murfield, J. E. (2010). A randomized controlled trial exploring the effect of music on agitated behaviours and anxiety in older people with dementia. *Aging and Mental Health*, *14*, 905-916. doi: 10.1080/13607861003713190

Clair, A. A., Mathews, R. M., & Kosloski, K. (2005). Assessment of active music participation as an indication of subsequent music making engagement for persons with

midstage dementia. *American Journal of Alzheimer's Disease & Other Dementias®*, *20*, 37-40. doi: 10.1177/153331750502000105

Curyto, K. J., Van Haitsma, K., & Vriesman, D. K. (2008). Direct observation of behavior: A review of current measures for use with older adults with dementia. *Research in Gerontological Nursing*, *1*, 52-76. doi: 10.3928/19404921-20080101-02

Deshmukh, S. R., Holmes, J., & Cardno, A. (2018). Art therapy for people with dementia. *Cochrane Database of Systematic Reviews*, *13*. doi: 10.1002/14651858.CD011073.pub2

Department of Health (2015). *Prime Minister's challenge on dementia 2020*. London, UK: Department of Health.

Dunne, R. (2002). *Dementia care planning: An identity-focused approach.* City, PA.: Venture Publishing.

Eggert, J., Dye, C. J., Vincent, E., Parker, V., Daily, S. B., Pham, H., ... & Roy, T. (2015). Effects of viewing a preferred nature image and hearing preferred music on engagement, agitation, and mental status in persons with dementia. *SAGE Open Medicine*, *3*, 2050312115602579. doi: 10.1177/2050312115602579

Engström, G., Marmstål Hammar, L., Williams, C., & Götell, E. (2011). The impact of singing in caring for a person with dementia: Single case analysis of video recorded sessions. *Music and Medicine, 3*, 95-101. doi: 10.1177/1943862110386048

Gaugler, J. E., Hobday, J. V., & Savik, K. (2013). The CARES® observational tool: A valid and reliable instrument to assess person-centered dementia care. *Geriatric Nursing*, *34*, 194-198. doi: 10.1016/j.gerinurse.2013.01.002

Gerdner, L. A. (2012). Individualized music for dementia: Evolution and application of evidence-based protocol. *World Journal of Psychiatry*, *2*, 26-32. doi: 10.5498/wjp.v2.i2.26

Greenwood, D., Loewenthal, D., & Rose, T. (2001). A relational approach to providing care for a person suffering from dementia. *Journal of Advanced Nursing*, *36*(4), 583-590. doi: 10.1046/j.1365-2648.2001.02011.x

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-based nursing*, 18, 66-67. doi: 10.1136/eb-2015-102129

Heck, R. H., & Thomas, S. L. (2000). *An introduction to multilevel modeling techniques*. Mahwah, NJ: Lawrence Erlbaum Associates.

Herrmann, N., & Gauthier, S. (2008). Diagnosis and treatment of dementia: Management of severe Alzheimer's disease. *Canadian Medical Association Journal*, *179*, 1279-1287. doi 10.1503/cmaj.070804

Holden, R. R. (2010). Face validity. *The corsini encyclopedia of psychology*, 1, 1-2. doi: 10.1002/9780470479216.corpsy0341

Hubbard, G., Cook, A., Tester, S., & Downs, M. (2002). Beyond words: Older people with dementia using and interpreting nonverbal behaviour. *Journal of Aging Studies*, *16*, 155-167. doi: 10.1016/S0890-4065(02)00041-5

Jones, C., Sung, B., & Moyle, W. (2015). Assessing engagement in people with dementia: A new approach to assessment using video analysis. *Archives of Psychiatric Nursing*, *29*, 377-382. doi: 10.1016/j.apnu.2015.06.019

Kawulich, B. (2012). Collecting data through observation. In C, Wagner., B, Kawulich., & M, Garner. (Eds.) *Doing social research: A global context*, (pp. 150-160). US, NY :McGraw-Hill Education

Kitwood, T. (1997). The experience of dementia. *Aging & Mental Health,1*, 13-22. doi: 10.1080/13607869757344

Kitwood, T. M. (1997). *Dementia reconsidered: The person comes first*. Buckingham, UK: Open University Press.

Koo T, Li M. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*,*15*, 155-63. doi: 10.1016/j.jcm.2016.02.012

Kolanowski, A. M., Van Haitsma, K., Meeks, S., & Litaker, M. (2014). Affect balance and relationship with well-being in nursing home residents with dementia. *American*

Journal of Alzheimer's Disease & Other Dementias, *29*, 457-462. doi: 10.1177/1533317513518657

Kumar, C. S., & Kuriakose, J. R. (2013). End-of-life care issues in advanced dementia. *Mental Health in Family Medicine*, *10*, 129. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3822659/

Lai, C. (2015). Person-centered dementia care. *The Encyclopedia of Adulthood and Aging*, 1-5. doi: 10.1002/9781118521373.wbeaa235

Lancaster G.A, Dodd S, Williamson P.R. (2004) Design and analysis of pilot studies: Recommendations for good practice. *Journal of Evaluation in Clinical Practice. 10*, 307-312 8. doi:10.1111/j.2002.384.doc.x

Lewis, J. R. (1993). Multipoint scales: Mean and median differences and observed significance levels. *International Journal of Human-Computer Interaction*, *5*, 383-392. doi: 10.1080/10447319309526075

Margallo-Lana, M., Swann, A., O'Brien, J., Fairbairn, A., Reichelt, K., Potkins, D., ... Ballard, C. (2001). Prevalence and pharmacological management of behavioural and psychological symptoms amongst dementia sufferers living in care environments. *International journal of geriatric psychiatry*, *16*, 39-44. doi: 10.1002/1099-1166

Maroulas, V., & Panagiotakos, D. B. (2010). *Sensitivity of health-related scales is a non-decreasing function of their classes*. Retrieved from https://arxiv.org/pdf/1005.3737.pdf

Martyr, A., Nelis, S. M., Quinn, C., Wu, Y. T., Lamont, R. A., Henderson, C., ... Morris, R. G. (2018). Living well with dementia: A systematic review and correlational meta-analysis of factors associated with quality of life, well-being and life satisfaction in people with dementia. *Psychological Medicine*, *48*, 2130-2139. doi: 10.1017/S0033291718000405

Medical Research Council. (2000, April). *A framework for the development and evaluation of RCTs for complex interventions to improve health.* Retrieved from https://mrc.ukri.org/documents/pdf/rcts-for-complex-interventions-to-improve-health/

Mental Health Foundation. (2011, October). *Spotlight on dementia care: A health foundation improvement report*. Retrieved from https://www.health.org.uk/sites/default/files/SpotlightOnDementiaCare_fullversion.pdf

Mitchell, G., & Agnelli, J. (2015). Person-centred care for people with dementia: Kitwood reconsidered. *Nursing Standard, 30*, 46. doi: 10.7748/ns.30.7.46.s47

Moore, A., Patterson, C., White, J., House, S. T., Riva, J. J., Nair, K., ... & McCann, D. (2012). Interprofessional and integrated care of the elderly in a family health team. *Canadian Family Physician*, 58, e436-e441. Retrieved from http://www.cfp.ca/content/cfp/58/8/e436.full.pdf

Music for Life. (July, 2014). *Music for life programme report: April 2013 – March 2014*. Retrieved from: https://wigmore-hall.org.uk/learning/23-music-for-life-report-2013-14/file.

Orrell, M., Yates, L., Leung, P., Kang, S., Hoare, Z., Whitaker, C., ... & Pearson, S. (2017). The impact of individual cognitive stimulation therapy (iCST) on cognition, quality of life, caregiver health, and family relationships in dementia: A randomised controlled trial. *PLoS Medicine*, *14*. doi: 10.1371/journal.pmed.1002269

Panagiotakos, D. B., & Yfantopoulos, J. N. (2011). Methodological issues in the quantitative assessment of quality of life. *Angiology*, *62*, 517-519. doi: 10.1177/0003319711403735

Raglio, A., Bellandi, D., Baiardi, P., Gianotti, M., Ubezio, M. C., Zanacchi, E., ... & Stramba-Badiale, M. (2015). Effect of active music therapy and individualized listening to music on dementia: A multicenter randomized controlled trial. *Journal of the American Geriatrics Society*, *63*, 1534-1539. doi: 10.1111/jgs.13558

Robert, P., Ferris, S., Gauthier, S., Ihl, R., Winblad, B., & Tennigkeit, F. (2010). Review of Alzheimer's disease scales: Is there a need for a new multi-domain scale for therapy evaluation in medical practice? *Alzheimer's Research & Therapy*, *2*, 24-37. doi: 10.1186/alzrt48

Rose, L. (1993). *Music for Life.* Retrieved from: https://wigmore-hall.org.uk/learning/music-for-life on 3/8/2019.

Sabat, S. (2006). Mind, meaning and personhood in dementia: the effects of positioning. In J, Hughes., S, Louw. and S, Sabat. (Eds.), *Dementia: Mind, meaning and the person*. Oxford: Oxford University Press.

Smit, D., De Lange, J., Willemse, B., Twisk, J., & Pot, A. M. (2016). Activity involvement and quality of life of people at different stages of dementia in long term care facilities. *Aging & Mental Health*, *20*, 100-109. doi: 10.1080/13607863.2015.1049116

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion (pp 227-239). In Levine, M.J,.& Moreland. L R., (Eds.) *Small groups.* City, N.Y.: Psychology Press.

Sung, H. C., & Chang, A. M. (2005). Use of preferred music to decrease agitated behaviours in older people with dementia: A review of the literature. *Journal of Clinical Nursing*, *14*, 1133-1140. doi: 10.1111/j.1365-2702.2005.01218.x

Thomas, R. Prohaska, T. R., Anderson, L. A., Binstock, R. H. (2012). *Public Health for an Aging Society*. Baltimore: The John Hopkins University Press.

Vasionytė, I., & Madison, G. (2013). Musical intervention for patients with dementia: A meta-analysis. *Journal of Clinical Nursing*, 22, 1203-1216. doi: 10.1111/jocn.12166

Zhang, S., Jiang, H., & Carroll, J. M. (2011). Integrating online and offline community through Facebook. *2011 International Conference on Collaboration Technologies and Systems (CTS)*, 569-578. doi: 10.1109/CTS.2011.5928738

Section C – Appendices

Appendix A. *VASE html source code*

```
<!DOCTYPE html>
<html>
<meta charset="utf-8">
<meta http-equiv="X-UA-Compatible" content="IE=edge">
<meta name="viewport" content="width=device-width, initial-scale=1">
<head><title>replace-tab-title</title></head>
<body>
<style type="text/css">
body{
  background-color: #b0e2d4;
  font-family: "Raleway", sans-serif;
}
.container {
    margin-right: auto;
    margin-left: auto;
    padding-left: 15px;
    padding-right: 15px;
    width:90%;
}
.center-block {
    display: block;
    margin-left: auto;
    margin-right: auto;
}
.site--panel{
    background: #fff;
    border-radius: 8px;
    -webkit-box-shadow: 0 0 40px -10px #000;
    box-shadow: 0 0 40px -10px #000;
    margin-top: 10px;
    /* margin: calc(0vh - 20px) auto; */
    padding: 10px 10px;
    /* max-width: calc(100vw - 40px); */
    -webkit-box-sizing: border-box;
    /*box-sizing: border-box;*/
    /* font-family: 'Montserrat',sans-serif; */
    position: relative;
    border-color: transparent;
 width:870px;

    margin-bottom: 22px;
}
.video-wrap{
  border: 1px solid #eeeeee;
   margin-right: auto;
    /*margin-left: auto;*/
/*    padding: 15px;*/
    width: 850px;

}
```

```
.slider-wrap{
  width:100%;
  padding:10px 0;
}
input[type="range"] {
    display: block;
    width: 100%;
}
.btn-results{
  border:1px solid #000;
  cursor: pointer;
}
.range-left{
  padding-right:5px;
  margin-top:-10px;
}
.range-right{
  padding-left:5px;
}
.btn-warning {
    color: #fff;
    background-color: #2ab27b;
    border-color: #259d6d;
}
.btn {
    display: inline-block;
    margin-bottom: 0;
    font-weight: normal;
    text-align: center;
    vertical-align: middle;
    -ms-touch-action: manipulation;
    touch-action: manipulation;
    cursor: pointer;
    background-image: none;
    border: 1px solid transparent;
    white-space: nowrap;
    padding: 6px 12px;
    font-size: 14px;
    line-height: 1.6;
    border-radius: 4px;
    -webkit-user-select: none;
    -moz-user-select: none;
    -ms-user-select: none;
    user-select: none;
}
.slider-range{
  margin:10px 0;
}
.slider-range .max { float:right }
.slider-range .min { float:left }
.pad-top-bot{
  margin:10px 0;
}
textarea#results {
  display:block;
  width: 100%;
  height: 120px;
  border: 2px solid #cccccc;
```

```
}
.heading{
  margin:10px 0;
}
</style>
<div class="container">

      <div class="panel panel-default site--panel">
         <div class="panel-heading clearfix"><h1 class="error">Music-1</h1> </div>
            <div class="video-wrap">
            <video id="myVideo" width="850" height="500" controls>
              <source src="file:set1.mp4" type="video/mp4" />
              <source src="file:set1.mp4" type="video/ogg" />
              Your browser does not support HTML5 video.
            </video>

         <div class="pad-top-bot">Playback position: <span id="txtPlaytime"></span></div>
         <div class="pad-top-bot">Rating : <span id="txtRating"></span></div>
         <div class="slider-range">
           <span class="min">1</span>
           <span class="max">7</span>
         </div>
         <div class="slider-wrap">
            <input type="range" min="1" max="7" value="0" list="ticks"/>

            <datalist id="ticks">
               <option>1</option>
               <option>2</option>
               <option>3</option>
               <option>4</option>
               <option>5</option>
               <option>6</option>
               <option>7</option>
            </datalist>
         </div>
<div class="btn btn-warning" onclick="showRes();">click me when done</div>
<!-- <div class="btn btn-warning" onclick="CopyToClipboard('results');">Show res (click me at the
end)</div> -->
<div class="heading">Results: Click anywhere in box below, Ctrl-a = select all, ctrl-v to copy, ctrl-c to
paste</div>
<textarea id="results" class="styled"></textarea>
 </div><!-- end video wrap -->
</div>
</div>
<script>
  window.Res = [];
  //find the vid div
  var vid = document.getElementById("myVideo");
  // Assign an ontimeupdate event to the video
  vid.ontimeupdate = function() {videoPlayTimeUpdate()};
  function videoPlayTimeUpdate() {
     // Display the current position of the video
     document.getElementById("txtPlaytime").innerHTML = vid.currentTime;
  }
  function getRes(){
    //dump the results out
    var str = 'playtime ,rating \r\n';
```

```
    window.Res.forEach(function(i){
          str += i + '\r\n';
    });

    return str;
  }
  function showRes(){

    document.getElementById("results").value = getRes();
  }
  //our slider
  var rng = document.querySelector("input");
  //listen function for slider
  var listener = function() {
    window.requestAnimationFrame(function() {
      window.Res.push(vid.currentTime +' , '+rng.value  );
      document.getElementById("txtRating").innerHTML = rng.value;
      // document.getElementById("results").innerHTML = rng.value + ' at ' + vid.currentTime;
    });
  };

rng.addEventListener("change", function() {
    listener();

  });
</script>
</body>
</html>
```

Appendix B. *Examples of categories and behavioural expressions of engagement*

	Category	Definition	Behaviour expression of engagement	Examples of the feedback (s)	No. of participants commenting on the categories
1	Facial Expression	Noticeable changes on the PWD's face during the intervention	*Mouth and lip movement*	"mouth moving, mumbling the song"	2
			Eyebrows movement (e.g. closed their eyes or raise their eyebrow)	"The person's eyebrows were raising when the music was playing"	1
			Facial changes (e.g. Neutral look, smile)	"I can see that the person (PWD) face looked different … like she was smiling) "The person face looked very neutral without much facial expression, but you can feel that she was enjoying the music, as if she was thinking about it."	6

2	Bodily Movement and verbal articulations	Large or small bodily movements and response during the intervention	*Large and subtitle bodily movement (e.g. Hands and feet taping, nodding, clapping, moving with music)*	"Hand clapping and feet tapping" "There was one person who tapped his hands on his lap"	6
			Verbally responding (e.g. singing, talking, mouth mumbling)	"One of the elderly was moving her mouth."	2
			Interacting with instruments (touching the instrument, playing with the instruments, making music)	"Hitting the African drum and the hand drum" "Playing with the drum stick"	3
3	Attention and awareness of activity	Being focus and attend to a stimulate that is in context with the intervention	*Attention to Stimulus (musician, other participants) undistracted eye contact*	"there's a lot of duplication. Because when they would said to you, "okay, this man has been accepted, could you please do a referral", and we are all using information on the same system. So we end up doing the same thing again" "We're having these discussions with psychology in the pod meeting they're quite often quite like in-depth, which is great, but then if they're accepted there and	4

				then in the meeting for psychology, the clinician, like the care coordinator then has to go away and type out the conversation. But we've already had the conversation with psychology, so could the referral not just be accepted there and then, without the paper part being done?"	
			Playing an instrument	"Playing and focusing on the instruments in front of her"	
			Moving along with music	"Her feet was tapping (along with the music)"	
4	Emotional response	Participant's "positive" and "negative" emotions in relations to the intervention	*Pleasure and enjoyment*	"I don't really know what's going on down here"	4
			At ease look (looks as if s/he was relaxed)	"She closed her eyes, but it looks like she was enjoying the music." "Looking up but thinking about things but she seemed relaxed"	2
			Sad or anxious look (appear agitated, e.g., eyes down casted like in moment of unhappiness; tapping his/her fingers as in people who are anxious)	"I can see that the elderly was sad…but it does not mean she was not enjoying the music right? Maybe it made her remember something" "But I suppose negative emotions like looked anxious and sad can mean that the person is (PWD) is engaging with the therapy (intervention) right ?)"	2

This has been removed from the electronic copy

Appendix D. *Information sheet for Family members*

Information about the Projects for Family Members

Hello. Our names are Amy Clare, Daniel Lai, Professor Paul Camic and Professor Seb Crutch. We are researchers at Canterbury Christ Church University and University College London; Amy and Daniel are completing their doctoral degrees in clinical psychology. Paul and Seb are the supervisors of the project; both are highly experienced researchers in dementia care and the arts. We would like to invite you to take part in a research project about the nature of verbal and non-verbal communication within a Music for Life group taking place at Jewish Care.

Before you decide, it is important that you understand why the research is being done and what it would involve for the person you are consenting for. You are welcome to tell others about the study if you wish.

Part 1 of this sheet tells you the purpose of this study and what will happen if the person take part.
Part 2 gives you more detailed information about how the study will be conducted.

Part 1

What is the purpose of the study?

The purpose of the research is to gain an understanding about the nature of communication within a Music for Life group. There are two parts to this research project. Both will use the same video recordings of Music for Life sessions. The first project will explore the nature of how people might communicate or express their emotions without using words, the nature of the interactions between all the people present in the group and how music may impact on communication with or without words. The second project of the study will be using the video recorded footage to help develop a video analysing tool to assess people with dementia's level of engagement in a Music for Life group. The development of the two will enable us to investigate the process that occurs within the Music for Life group, and potentially identify factors which enable people with dementia to engage.

Why am I being provided with this information?

You are being provided this information because a person close to you will be taking part in the Music for Life group at Jewish Care. However, the person taking part may find it very difficult or not be able to give their own consent to being involved with the study. Usually, when this happens an individual that is close to the person who cannot give consent is asked on their behalf. The group will be video recorded and sections of these recordings will be watched as part of the study.

Your family member will also be asked to wear an Empatica E4 wristband for up to 3 of the sessions. The wristband looks like a watch and it measures the physical responses of heart rate, bodily movement and skin conductance. It feels no different than wearing a watch and causes no discomfort. If your family member wants to remove the wristband at any time after it is fitted, we will remove it as soon as they have indicated this to us verbally or non-verbally.

Does the person have to take part?
No, nobody has to take part. It is entirely up to you to decide whether the person is involved in the research. If you agree for them to take part, we will then ask you to sign a consent form. If you do not want them to take part in the research, this would not affect their ability to take part in the Music for Life group or any other activity at Jewish Care.

What are the possible disadvantages and risks of taking part?
There are no known disadvantages or risks for taking part in the Music for Life group. The video recording device is smaller than a tennis ball and is able to capture a 360 view of the group without any interference in the enjoyment and participation by group members.

What are the possible benefits of taking part?
The results of this study will be used to develop a better understanding of how people might express themselves with or without words. It will highlight the importance of noticing, valuing and responding to any forms of communication within a group setting, in order to improve wellbeing and relationships for people with more severe levels of impairment in dementia. It will also look at how music, singing and the musicians influence communication.

Will taking part in this study be kept confidential?
Each person's confidentiality will be safeguarded during and after the study. No participant in the Music for Life group will be identified. The video recordings will be kept on a password protected computer that only the researchers have access to. The video file will also be encrypted file, which assures its safety and confidentiality; we will not use any form of internet communication to transfer video files.

Who has approved the study?

This study has been approved by an independent research review panel at the Salomons Centre for Applied Psychology, Canterbury Christ Church University. It has also been reviewed and given approval by a Canterbury Christ Church University Research Ethics Committee on the 11th Aug 2017

Who is organising and funding the research?
Canterbury Christ Church University and the Created Out of Mind Hub at the Wellcome Collection are funding the research. Jewish Care and Wigmore Hall are helping organise the Music for Life groups and are partners in this research.

Part 2

What will happen if I don't want to the person to carry on with the study?
If you change your mind about the person participating in the research, but still want them to come to the groups, you can simply let Jewish Care know at the beginning of the group. You do not have to give a reason.

What if there is a problem?
If there is a problem you can ask one of the Jewish Care staff to help you or to give you more information.

Concerns and Complaints
If you have any *concerns* or questions about the research, please contact Professor Paul Camic, Salomons Centre for Applied Psychology, Canterbury Christ Church University, paul.camic@canterbury.ac.uk. You can also leave a message on Paul's telephone at 03330 117 114. He will get back to you as soon as possible.

If you would like to make a *complaint* about any aspect of the research please contact: Professor Margie Callanan, Chair, university ethics panel at Salomons Centre for Applied Psychology, Canterbury Christ Church University at margie.callanan@canterbury.ac.uk or telephone 03330 117 094.

What will happen to the results of the research study?
We will send you a brief review of the study when it is completed and the data is analysed, letting you know what we have learnt from it. Please inform Professor Camic if you do not wish to receive this. The results of the study will also be part of the doctoral dissertations of Amy Clare and Daniel Lai. It is hoped that the results of the research will also be published in journal articles.

Further information and contact details
If you would like to speak to the research team about the study or have questions, please contact Professor Paul Camic at paul.camic@canterbury.ac.uk If you provide a telephone number he can also call or email you.

Thank you

Appendix E. *Consent form for family members*

Consent Form for family members (page 1 of 2)

Participant ID: _____

Title of Project: Understanding communication in a Music for Life group

Name of Researchers: Amy Clare, Daniel Lai, Professor Paul Camic and Professor Seb Crutch.

Please initial each box if you agree

1. I confirm that I have read and understand the information sheet for the above research. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that the Music for Life group session will be video recorded for the purpose of this research. ☐

3. I also understand the video recordings may be used in possible future research and for educational purposes. ☐

4. I also understand that my family member will be asked to wear the Empatica wristband for one or two music sessions. I agree to them wearing it. ☐

5. If your relative had been able to give consent for this, would they have agreed to participate and do you think this is something they would have wanted? ☐

6. I understand that anonymous data from this project will be available to Canterbury Christ Church University and University College London researchers; results from this research will be submitted as part of two doctoral theses and for journal publications, and that information from the study may be used in future research projects. ☐

7. I understand that anonymous data from this project will be available to Canterbury Christ Church University and University College London researchers; results from this research will be submitted as part of two doctoral theses and for journal publications, and that information from the study may be used in future research projects. ☐

8. I give informed consent for the participant to take part in this present project and acknowledge that the participation is voluntary and that they are free to withdraw at any time without giving any reason. ☐

Name of Participant: _____

Name of Person giving consent: _____

Signature: _____

Date: _____

Appendix F. *Information sheet for raters*

Information about the Project for Volunteer Raters

Hello. Our names are Daniel Lai, Professor Paul Camic and Professor Seb Crutch. We are researchers at Canterbury Christ Church University and University College London; Daniel is completing his doctoral degree in clinical psychology and the lead investigator. Paul and Seb are the supervisors of the project; both are highly experienced researchers in dementia care and the arts.

We would like to invite you to take part in a research project about the nature of verbal and non-verbal communication within a music group for people with severe dementia.

Before you decide, it is important that you understand why the research is being done.

What is the purpose of the study?
The purpose of the study is to develop an observation tool to better understand the process of engagement for people with dementia who take part in a music-based group, using video observation technology. The development of the observation tool will enable us to understand the processes that occur within the music-based groups in dementia care, and potentially identify factors which help enable people with dementia to better engage.

Do I have to take part?
No, nobody has to take part. It is entirely up to you to decide whether you would like to be involved in the research. If you agree to take part, we will then ask you to sign a consent form. If you do not want them to take part in the research, you may decline at any time.

What you will be asked to do in the research
We are inviting you to view several short video clips of people with dementia taking part in a music group in a care home. There will be a brief introduction to how to use the rating system we are trying out. This will be followed by watching several brief video segments and rating what you observe. There are no right or wrong answers. Your involvement will take up to about 60 minutes and as a thank you for helping out, we will give you a shopping voucher worth £10.00.

What are the possible disadvantages and risks of taking part?
There are no known or anticipated risks in viewing video footage of music-based groups of people with dementia. It is possible, but unlikely, that you may have an emotional response when viewing people with severe dementia. If you should find any of the video upsetting, you can take a break and then resume, stop for now and come back at a different time, or decide to no longer participate in the research.

What are the possible benefits of taking part?
The results of this study will be used to develop a tool that potentially allows us to have a better understanding of how people with severe dementia might engage in a music group. Taking part will also enable you to witness the effect of a music group for people with severe dementia, which you may find interesting and beneficial to your own understanding of dementia and music.

Will my taking part in this study be kept confidential?
Each person's confidentiality will be safeguarded during and after the study. No participant in the research will be identified. Your responses as a rater will be kept on a password-protected computer that only the researchers have access to, which assures its safety and confidentiality.

Who has approved the study?
This study has been approved by an independent research review panel at the Salomons Centre for Applied Psychology, Canterbury Christ Church University. It has also been reviewed and given approval by a Canterbury Christ Church University Research Ethics Committee on the 11th Aug 2017

Who is organising and funding the research?
Canterbury Christ Church University and the Created Out of Mind Hub at the Wellcome Collection are funding the research. Jewish Care and Wigmore Hall are helping organise the Music for Life groups and are partners in this research.

What will happen if I don't want to carry on with the study?
If you change your mind about participating in the research or want to withdraw at any time, contact Daniel Lai at (d.l.lai234@canterbury.ac.uk) to ask to withdraw. No reason will be need to be given to withdraw from the research.

What if there is a problem?
If there is a problem you can contact Mr Daniel Lai via d.l.lai234@canterbury.ac.uk or alternative contact the unviersty research number on 03330 117 094.

To make a complaint about the research

If you would like to make a *complaint* about any aspect of the research please contact Professor Margie Callanan, Chair, University ethics panel at Salomons Centre for

Applied Psychology, Canterbury Christ Church University at margie.callanan@canterbury.ac.uk or telephone 03330 117 094.

What will happen to the results of the research study?
If you would like us to do so, we can send you a brief review of the study when it is completed and the data is analysed, letting you know what we have learnt from it. Please inform Daniel Lai at (add email) .The results of the study will also be part of the doctoral dissertations of Daniel Lai. It is hoped that the results of the research will also be published in journal articles.

Further information and contact details
If you have further questions, please email Daniel Lai and he can call or email you back. Please leave your telephone number and times to call.

Thank you

Appendix G. *Consent form for raters*

Attachment 3: Consent Form for Volunteer raters (professional and people)

Consent Form (page 1 of 2)

Participant ID: _____

Title of Project: Understanding communication in a Music for Life group

Name of Researchers: Amy Clare, Daniel Lai, Professor Paul Camic and Professor Seb Crutch.

Please initial each box if you agree

1. I confirm that I have read and understood the information sheet for the above research. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that the video I will be watching features vulnerable individuals. I agree never to disclose the name or identifying information about any of the participants involved in the video. ☐

3. I understand that anonymous data from this project will be available to Canterbury Christ Church University and University College London researchers. ☐

4. I understand that results from this research will be submitted as part of a doctoral thesis and for journal publications, and that information from the study may be used in future research projects and educational purposes. ☐

5. I give informed consent to take part in this research project and acknowledge that my participation is voluntary and I am free to withdraw at any time without giving any reason. ☐

Name of Participant: _____

Signature: _____

Date: ____

Appendix H. *Final version of the VASE protocol*

The video clips that you will see are taken from a music session for people with a dementia. The session consists of musicians, care staff and people with a dementia improvising music together.

When watching the video clips, please rate the level of engagement you observe during the clip. You will be shown video clips that focus on different people with dementia. I will demonstrate how to use the rating scale within the tablet before you begin.

There are no right or wrong answers. We would like you to carefully observe the video clips whilst rating the level of engagement.

Engagement might be perceived in many different ways, ranging from bodily movements through to someone sitting still and simply listening. We have asked some professionals and the general public about what they think is important when considering engagement in people with dementia. Below are some of the themes we drew from them, the themes might be helpful for you to consider when rating the session.

1. Facial Expression and responses.
   Examples:
   - Mouth and lip movement
   - Eyebrows movement
   - Neutral look

2. Bodily Movement and behavioral responses:
   Examples:
   - Large and small bodily movements
   - Verbal responses such as singing or movements of the mouth
   - Playing or exploring an instrument

3. Attention and awareness of the activity:
   Examples:
   - Watching the musicians or other residents
   - Playing an instrument
   - Sense of awareness of the activity
   - Eye contact with others

4. Emotional responses:
   Examples:
   - Pleasure and enjoyment or displeasure
   - Excitement
   - Anxiety (Agitation)
   - Sadness
   - Neutral
   - Being 'moved'

Appendix I: *Table A1. Distribution of rating over time analyzed by raters' characteristics*

Seconds	Overall			PL			Active			Professionals			Non-professionals		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
6.0	3.0	1.94	3.0	2.4	1.78	2.0	3.3	1.95	3.0	3.0	1.97	2.0	3.4	1.78	3.0
7.0	3.1	1.95	3.0	2.4	1.75	2.0	3.3	1.96	3.0	3.0	1.97	2.0	3.5	1.81	3.0
8.0	3.1	1.97	3.0	2.4	1.76	2.0	3.4	1.97	3.0	3.0	1.99	3.0	3.5	1.83	3.0
9.0	3.2	1.97	3.0	2.4	1.73	2.0	3.4	1.97	3.0	3.1	1.99	3.0	3.5	1.81	3.0
10.0	3.2	1.98	3.0	2.4	1.71	2.0	3.5	1.99	3.0	3.1	2.00	3.0	3.6	1.85	3.0
11.0	3.3	1.98	3.0	2.3	1.71	2.0	3.6	1.98	3.0	3.2	2.01	3.0	3.6	1.84	4.0
12.0	3.3	2.00	3.0	2.3	1.72	2.0	3.6	1.99	3.0	3.2	2.03	3.0	3.6	1.83	4.0
13.0	3.4	2.01	3.0	2.3	1.72	2.0	3.7	1.98	4.0	3.3	2.05	3.0	3.6	1.81	4.0
14.0	3.4	2.01	3.0	2.3	1.72	2.0	3.7	1.99	4.0	3.3	2.05	3.0	3.7	1.83	4.0
15.0	3.4	2.01	3.0	2.4	1.71	2.0	3.8	1.98	4.0	3.3	2.04	3.0	3.7	1.85	4.0
16.0	3.4	2.02	3.0	2.4	1.72	2.0	3.8	1.99	4.0	3.4	2.05	3.0	3.7	1.85	4.0
17.0	3.4	2.02	3.0	2.4	1.71	2.0	3.8	2.00	4.0	3.4	2.06	3.0	3.7	1.83	4.0
18.0	3.5	2.03	3.0	2.4	1.70	2.0	3.8	2.00	4.0	3.4	2.07	3.0	3.7	1.84	4.0
19.0	3.5	2.03	3.0	2.4	1.68	2.0	3.8	2.01	4.0	3.4	2.07	3.0	3.8	1.85	4.0
20.0	3.5	2.03	3.0	2.4	1.67	2.0	3.8	2.01	4.0	3.4	2.07	3.0	3.8	1.85	4.0
21.0	3.5	2.04	3.0	2.4	1.68	2.0	3.9	2.02	4.0	3.4	2.08	3.0	3.8	1.84	4.0
22.0	3.5	2.05	3.0	2.4	1.69	2.0	3.9	2.02	4.0	3.5	2.08	3.0	3.8	1.85	4.0
23.0	3.5	2.04	3.0	2.4	1.70	2.0	3.9	2.00	4.0	3.5	2.08	3.0	3.8	1.82	4.0
24.0	3.5	2.03	3.0	2.4	1.70	2.0	3.9	2.00	4.0	3.5	2.08	3.0	3.8	1.81	4.0
25.0	3.6	2.03	3.0	2.5	1.69	2.0	3.9	2.01	4.0	3.5	2.08	3.0	3.8	1.80	4.0
26.0	3.6	2.04	3.0	2.5	1.72	2.0	4.0	2.01	4.0	3.5	2.09	3.0	3.8	1.81	4.0
27.0	3.6	2.05	3.0	2.6	1.75	2.0	4.0	2.01	4.0	3.6	2.09	3.0	3.8	1.83	4.0
28.0	3.6	2.04	3.0	2.6	1.75	2.0	4.0	2.00	4.0	3.6	2.09	3.0	3.8	1.83	4.0
29.0	3.7	2.05	3.0	2.5	1.75	2.0	4.0	2.00	4.0	3.6	2.09	3.0	3.9	1.82	4.0
30.0	3.7	2.04	3.0	2.6	1.76	2.0	4.0	2.00	4.0	3.6	2.09	3.0	3.9	1.81	4.0
Average	3.4	2.02	3.0	2.4	1.72	2.0	3.8	2.01	4.0	3.4	2.06	3.0	3.7	1.83	4.0

Table A1. *Distribution of rating over time analyzed by raters' characteristics (Cont'd)*

Seconds	Male			Female			M1			M2			M3		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
6.0	3.1	1.86	3.0	3.0	1.98	2.0	3.3	2.03	3.0	3.1	2.03	3.0	2.7	1.70	2.0
7.0	3.1	1.87	3.0	3.1	1.99	3.0	3.4	2.02	3.0	3.1	2.05	3.0	2.8	1.72	2.0
8.0	3.2	1.88	3.0	3.1	2.00	3.0	3.4	2.01	3.0	3.2	2.07	3.0	2.8	1.75	2.0
9.0	3.2	1.88	3.0	3.1	2.01	3.0	3.4	2.00	3.0	3.2	2.08	3.0	2.9	1.76	2.0
10.0	3.3	1.90	3.0	3.2	2.02	3.0	3.5	2.01	3.0	3.2	2.11	3.0	2.9	1.76	3.0
11.0	3.3	1.89	3.0	3.2	2.03	3.0	3.5	2.01	3.0	3.3	2.12	3.0	2.9	1.76	3.0
12.0	3.4	1.91	3.0	3.3	2.05	3.0	3.5	2.02	3.0	3.4	2.14	3.0	3.0	1.78	3.0
13.0	3.4	1.91	3.0	3.3	2.06	3.0	3.5	2.02	3.0	3.4	2.16	3.0	3.1	1.80	3.0
14.0	3.4	1.92	3.0	3.4	2.06	3.0	3.6	2.04	3.0	3.5	2.16	3.0	3.1	1.79	3.0
15.0	3.5	1.93	3.0	3.4	2.05	3.0	3.6	2.04	3.0	3.5	2.16	3.0	3.2	1.79	3.0
16.0	3.5	1.94	3.0	3.4	2.06	3.0	3.6	2.03	3.0	3.5	2.17	3.0	3.2	1.82	3.0
17.0	3.5	1.93	3.0	3.4	2.07	3.0	3.6	2.03	3.0	3.5	2.17	3.0	3.2	1.82	3.0
18.0	3.5	1.94	3.0	3.4	2.07	3.0	3.6	2.05	3.0	3.5	2.17	3.0	3.2	1.83	3.0
19.0	3.5	1.94	3.0	3.4	2.07	3.0	3.6	2.06	3.0	3.5	2.16	3.0	3.3	1.83	3.0
20.0	3.5	1.95	3.0	3.5	2.07	3.0	3.6	2.07	3.0	3.5	2.16	3.0	3.3	1.83	3.0
21.0	3.5	1.95	3.0	3.5	2.09	3.0	3.7	2.07	3.0	3.5	2.17	3.0	3.3	1.85	3.0
22.0	3.5	1.95	3.0	3.5	2.09	3.0	3.7	2.07	3.0	3.5	2.17	3.0	3.3	1.86	3.0
23.0	3.5	1.94	3.0	3.5	2.08	3.0	3.7	2.06	3.0	3.6	2.17	3.0	3.3	1.84	3.0
24.0	3.6	1.95	3.0	3.5	2.07	3.0	3.7	2.06	3.0	3.6	2.16	3.0	3.3	1.83	3.0
25.0	3.6	1.95	3.0	3.5	2.07	3.0	3.7	2.06	3.0	3.6	2.16	3.0	3.4	1.84	3.0
26.0	3.6	1.96	3.0	3.6	2.08	3.0	3.7	2.06	3.0	3.6	2.17	3.0	3.4	1.86	3.0
27.0	3.6	1.95	3.0	3.6	2.09	3.0	3.8	2.07	3.5	3.6	2.16	3.0	3.4	1.88	3.0
28.0	3.6	1.94	3.0	3.7	2.09	3.0	3.8	2.07	4.0	3.7	2.15	3.0	3.5	1.88	3.0
29.0	3.6	1.95	3.0	3.7	2.09	3.0	3.8	2.06	4.0	3.7	2.16	3.0	3.5	1.89	3.0
30.0	3.6	1.94	3.0	3.7	2.09	3.0	3.8	2.05	4.0	3.7	2.16	3.0	3.5	1.88	3.0
Average	3.5	1.93	3.0	3.4	2.07	3.0	3.6	2.05	3.0	3.5	2.15	3.0	3.2	1.83	3.0

Table A1. *Distribution of rating over time analyzed by raters' characteristics (Cont'd)*

Seconds	PL			Session 1			Session 2			Session 3			Session 4		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
6.0	2.4	1.78	2.0	3.1	1.95	3.0	3.2	1.80	3.0	2.9	1.74	3.0	3.1	1.86	3.0
7.0	2.4	1.75	2.0	3.1	1.94	3.0	3.2	1.82	3.0	2.8	1.68	3.0	3.3	1.86	3.0
8.0	2.4	1.76	2.0	3.2	1.99	3.0	3.3	1.83	3.0	2.8	1.65	3.0	3.3	1.84	3.0
9.0	2.4	1.73	2.0	3.3	2.00	3.0	3.4	1.82	3.0	2.8	1.63	3.0	3.2	1.78	3.0
10.0	2.4	1.71	2.0	3.3	2.01	3.0	3.5	1.86	3.5	2.8	1.62	3.0	3.3	1.74	3.0
11.0	2.3	1.71	2.0	3.4	2.04	3.0	3.7	1.87	4.0	2.8	1.67	3.0	3.3	1.70	3.0
12.0	2.3	1.72	2.0	3.6	2.16	3.0	3.8	1.85	4.0	3.0	1.70	3.0	3.3	1.72	3.0
13.0	2.3	1.72	2.0	3.7	2.18	4.0	4.0	1.85	4.0	3.0	1.69	3.0	3.4	1.75	3.0
14.0	2.3	1.72	2.0	3.8	2.19	4.0	4.1	1.84	4.0	3.0	1.71	3.0	3.4	1.71	3.0
15.0	2.4	1.71	2.0	3.8	2.25	4.0	4.2	1.81	4.0	3.0	1.70	3.0	3.3	1.72	3.0
16.0	2.4	1.72	2.0	3.9	2.31	4.0	4.3	1.81	4.0	3.1	1.68	3.0	3.3	1.72	3.0
17.0	2.4	1.71	2.0	3.9	2.31	4.0	4.4	1.84	5.0	3.0	1.66	3.0	3.3	1.71	3.0
18.0	2.4	1.70	2.0	3.9	2.33	4.0	4.5	1.83	5.0	3.0	1.66	3.0	3.3	1.73	3.0
19.0	2.4	1.68	2.0	3.9	2.34	4.0	4.6	1.84	5.0	3.1	1.68	3.0	3.3	1.74	3.0
20.0	2.4	1.67	2.0	4.0	2.35	4.0	4.6	1.81	5.0	3.1	1.74	3.0	3.2	1.72	3.0
21.0	2.4	1.68	2.0	4.0	2.38	4.0	4.7	1.80	5.0	3.2	1.74	3.0	3.3	1.74	3.0
22.0	2.4	1.69	2.0	4.1	2.39	4.0	4.8	1.81	5.0	3.2	1.76	3.0	3.3	1.75	3.0
23.0	2.4	1.70	2.0	4.1	2.40	4.0	4.8	1.82	5.0	3.2	1.74	3.0	3.3	1.76	3.0
24.0	2.4	1.70	2.0	4.1	2.40	4.0	4.9	1.82	5.0	3.3	1.67	3.0	3.3	1.75	3.0
25.0	2.5	1.69	2.0	4.1	2.40	4.0	5.0	1.82	5.0	3.4	1.72	3.0	3.3	1.76	3.0
26.0	2.5	1.72	2.0	4.1	2.41	4.0	5.1	1.80	5.0	3.4	1.72	3.0	3.3	1.79	3.0
27.0	2.6	1.75	2.0	4.1	2.40	4.0	5.1	1.78	5.0	3.5	1.71	3.0	3.3	1.81	3.0
28.0	2.6	1.75	2.0	4.1	2.40	4.0	5.1	1.73	5.0	3.5	1.72	3.0	3.4	1.81	3.0
29.0	2.5	1.75	2.0	4.1	2.38	4.0	5.1	1.73	5.0	3.5	1.72	3.0	3.5	1.83	3.0
30.0	2.6	1.76	2.0	4.1	2.38	4.0	5.1	1.72	5.0	3.6	1.71	3.0	3.5	1.83	3.0
Average	2.4	1.72	2.0	3.8	2.28	4.0	4.3	1.92	5.0	3.1	1.71	3.0	3.3	1.76	3.0

Table A1 *Distribution of rating over time analyzed by raters' characteristics (Cont'd)*

Seconds	Session 5			Session 6			Session 7			Session 8		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
6.0	4.0	2.30	4.0	3.3	1.99	3.0	3.3	1.83	3.0	3.0	1.82	3.0
7.0	4.1	2.35	4.0	3.4	2.04	3.0	3.5	1.86	3.0	3.0	1.77	3.0
8.0	4.1	2.35	4.0	3.4	2.08	3.0	3.6	1.84	4.0	3.1	1.78	3.0
9.0	4.2	2.40	4.0	3.5	2.10	3.0	3.6	1.84	4.0	3.2	1.76	3.0
10.0	4.2	2.39	5.0	3.6	2.12	3.0	3.7	1.88	4.0	3.3	1.81	3.0
11.0	4.2	2.37	5.0	3.6	2.09	4.0	3.8	1.85	4.0	3.4	1.79	3.0
12.0	4.2	2.34	4.0	3.7	2.09	3.0	3.8	1.86	4.0	3.5	1.79	3.0
13.0	4.2	2.33	4.0	3.7	2.10	3.0	3.9	1.86	4.0	3.6	1.76	3.0
14.0	4.2	2.32	4.0	3.8	2.13	4.0	3.9	1.85	4.0	3.7	1.73	3.0
15.0	4.1	2.30	4.0	3.8	2.14	4.0	4.0	1.82	4.0	3.7	1.73	3.0
16.0	4.1	2.29	4.0	3.8	2.15	4.0	4.0	1.80	4.0	3.7	1.76	3.0
17.0	4.1	2.30	4.0	3.8	2.16	4.0	4.0	1.80	4.0	3.7	1.76	3.0
18.0	4.1	2.32	4.0	3.9	2.14	4.0	3.9	1.80	4.0	3.8	1.77	4.0
19.0	4.0	2.31	4.0	3.9	2.14	4.0	4.0	1.79	4.0	3.8	1.75	3.0
20.0	4.0	2.34	4.0	3.9	2.15	4.0	4.0	1.75	4.0	3.8	1.76	4.0
21.0	4.1	2.35	4.0	4.0	2.12	4.0	4.0	1.73	4.0	3.8	1.75	4.0
22.0	4.0	2.34	4.0	4.0	2.10	4.0	4.0	1.70	4.0	3.9	1.76	4.0
23.0	4.0	2.28	4.0	4.0	2.10	4.0	4.0	1.69	4.0	3.8	1.73	4.0
24.0	4.0	2.30	4.0	3.9	2.10	4.0	4.0	1.72	4.0	3.9	1.71	4.0
25.0	3.9	2.29	4.0	3.9	2.07	4.0	4.0	1.73	4.0	3.9	1.75	4.0
26.0	3.9	2.31	4.0	4.0	2.04	4.0	4.0	1.73	4.0	3.9	1.75	4.0
27.0	3.9	2.32	4.0	4.0	2.03	4.0	4.0	1.77	4.0	4.0	1.76	4.0
28.0	3.9	2.34	4.0	4.0	1.99	4.0	4.1	1.79	4.0	4.0	1.73	4.0
29.0	3.9	2.31	4.0	4.0	1.97	4.0	4.1	1.84	4.0	4.0	1.74	4.0
30.0	4.0	2.30	4.0	4.1	1.96	4.0	4.1	1.84	4.0	3.9	1.76	4.0
Average	4.1	2.32	4.0	3.8	2.09	4.0	3.9	1.8	4.0	3.7	1.78	3.0

Table A1 *Distribution of rating over time analyzed by raters' characteristics (Cont'd)*

Seconds	Without PWD Experience			With PWD Experience			Singgroup=0			Singgroup=1		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
6.0	3.2	2.09	2.5	2.9	1.82	3.0	3.2	1.96	3.0	2.9	1.92	2.0
7.0	3.2	2.08	3.0	3.0	1.85	3.0	3.2	1.97	3.0	3.0	1.93	2.5
8.0	3.3	2.08	3.0	3.0	1.87	3.0	3.3	1.97	3.0	3.0	1.95	3.0
9.0	3.3	2.09	3.0	3.0	1.88	3.0	3.3	1.96	3.0	3.1	1.96	3.0
10.0	3.4	2.09	3.0	3.1	1.90	3.0	3.3	1.98	3.0	3.1	1.98	3.0
11.0	3.4	2.10	3.0	3.1	1.90	3.0	3.4	1.97	3.0	3.2	1.99	3.0
12.0	3.5	2.12	3.0	3.2	1.91	3.0	3.4	1.97	3.0	3.2	2.01	3.0
13.0	3.5	2.14	3.0	3.2	1.91	3.0	3.5	1.98	3.0	3.3	2.02	3.0
14.0	3.6	2.15	3.0	3.3	1.91	3.0	3.5	1.97	3.0	3.3	2.03	3.0
15.0	3.6	2.14	3.0	3.3	1.91	3.0	3.5	1.97	3.0	3.3	2.03	3.0
16.0	3.6	2.15	3.0	3.3	1.92	3.0	3.5	1.98	3.0	3.3	2.03	3.0
17.0	3.6	2.14	3.0	3.3	1.93	3.0	3.5	1.99	4.0	3.4	2.03	3.0
18.0	3.6	2.14	3.0	3.3	1.94	3.0	3.5	1.99	3.0	3.4	2.04	3.0
19.0	3.6	2.14	3.0	3.3	1.94	3.0	3.6	1.99	3.0	3.4	2.04	3.0
20.0	3.6	2.15	3.0	3.3	1.94	3.0	3.6	1.98	3.0	3.4	2.05	3.0
21.0	3.6	2.15	3.0	3.4	1.95	3.0	3.6	1.99	3.0	3.4	2.06	3.0
22.0	3.7	2.15	3.0	3.4	1.96	3.0	3.6	2.00	3.0	3.4	2.06	3.0
23.0	3.7	2.15	3.0	3.4	1.94	3.0	3.6	1.98	3.0	3.4	2.06	3.0
24.0	3.7	2.15	3.0	3.4	1.93	3.0	3.6	1.96	3.0	3.5	2.06	3.0
25.0	3.7	2.15	3.0	3.4	1.93	3.0	3.6	1.95	3.0	3.5	2.07	3.0
26.0	3.7	2.15	3.0	3.4	1.95	3.0	3.6	1.98	4.0	3.5	2.06	3.0
27.0	3.8	2.15	3.0	3.5	1.95	3.0	3.6	1.98	4.0	3.6	2.07	3.0
28.0	3.8	2.15	3.0	3.5	1.95	3.0	3.7	1.99	4.0	3.6	2.06	3.0
29.0	3.8	2.15	3.0	3.5	1.95	3.0	3.7	2.00	4.0	3.6	2.05	3.0
30.0	3.8	2.15	3.0	3.5	1.95	3.0	3.7	2.00	4.0	3.6	2.05	3.0
Average	3.6	2.14	3.0	3.3	1.93	3.0	3.5	1.98	3.0	3.3	2.03	3.0

Table A1 *Distribution of rating over time analyzed by raters' characteristics (Cont'd)*

Seconds	Groupresent=0			Grouppresent=1			Without Experience in Playing Instrument			Experience in Playing Instrument		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
6.0	3.0	1.92	2.0	3.2	1.98	3.0	2.9	1.90	2.0	3.2	1.97	3.0
7.0	3.0	1.93	3.0	3.3	1.99	3.0	2.9	1.92	2.0	3.2	1.96	3.0
8.0	3.1	1.95	3.0	3.3	1.99	3.0	3.0	1.94	2.0	3.3	1.97	3.0
9.0	3.1	1.95	3.0	3.3	2.01	3.0	3.0	1.94	2.0	3.3	1.98	3.0
10.0	3.2	1.97	3.0	3.3	2.00	3.0	3.0	1.97	3.0	3.4	1.99	3.0
11.0	3.2	1.98	3.0	3.3	2.00	3.0	3.1	1.97	3.0	3.4	1.99	3.0
12.0	3.3	1.99	3.0	3.4	2.03	3.0	3.1	1.98	3.0	3.5	2.01	3.0
13.0	3.3	1.99	3.0	3.4	2.05	3.0	3.2	1.98	3.0	3.5	2.02	3.0
14.0	3.4	2.00	3.0	3.5	2.05	3.0	3.2	1.99	3.0	3.6	2.03	3.0
15.0	3.4	1.99	3.0	3.5	2.05	3.0	3.2	1.99	3.0	3.6	2.02	3.0
16.0	3.4	2.00	3.0	3.5	2.06	3.0	3.2	1.99	3.0	3.6	2.03	3.0
17.0	3.4	2.00	3.0	3.5	2.08	3.0	3.3	2.00	3.0	3.6	2.02	3.0
18.0	3.4	2.00	3.0	3.5	2.09	3.0	3.3	2.00	3.0	3.6	2.03	3.0
19.0	3.4	2.00	3.0	3.5	2.09	3.0	3.3	2.00	3.0	3.7	2.02	3.0
20.0	3.4	2.00	3.0	3.6	2.10	3.0	3.3	2.00	3.0	3.7	2.03	3.0
21.0	3.5	2.01	3.0	3.6	2.11	3.0	3.3	2.01	3.0	3.7	2.04	3.0
22.0	3.5	2.02	3.0	3.6	2.09	3.0	3.3	2.03	3.0	3.7	2.03	3.0
23.0	3.5	2.01	3.0	3.6	2.09	3.0	3.3	2.01	3.0	3.7	2.03	3.0
24.0	3.5	2.00	3.0	3.7	2.10	3.0	3.3	1.99	3.0	3.7	2.04	3.5
25.0	3.5	2.00	3.0	3.7	2.11	3.0	3.3	1.99	3.0	3.7	2.04	4.0
26.0	3.5	2.01	3.0	3.7	2.10	3.0	3.4	2.00	3.0	3.8	2.05	4.0
27.0	3.5	2.02	3.0	3.7	2.09	3.0	3.4	2.01	3.0	3.8	2.05	4.0
28.0	3.6	2.02	3.0	3.7	2.07	4.0	3.4	2.01	3.0	3.8	2.04	4.0
29.0	3.6	2.03	3.0	3.7	2.05	4.0	3.4	2.02	3.0	3.8	2.04	4.0
30.0	3.6	2.03	3.0	3.7	2.04	4.0	3.5	2.01	3.0	3.8	2.04	4.0
Average	3.4	2.00	3.0	3.5	2.06	3.0	3.2	1.99	3.0	3.6	2.03	3.0

Appendix J. *Table A2. Spearman's correlation coefficient among non-professionals*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
A1_M1 (1)	1.000															
A5_M1 (2)	0.097**	1.000														
A8_M1 (3)	0.564**	0.033	1.000													
AS4_M1 (4)	0.647**	0.016	0.717**	1.000												
CCIL_3_M1 (5)	0.803**	0.107**	0.683**	0.711**	1.000											
CCIL_5_M1 (6)	0.715**	0.111**	0.632**	0.673**	0.819**	1.000										
CCIL_7_M1 (7)	0.662**	0.126**	0.629**	0.580**	0.768**	0.724**	1.000									
DRC5_M1 (8)	0.668**	-0.030	0.565**	0.584**	0.685**	0.591**	0.647**	1.000								
DRC8_M1 (9)	0.546**	0.042	0.478**	0.472**	0.687**	0.591**	0.656**	0.636**	1.000							
DRC9_M1 (10)	0.565**	0.021	0.559**	0.625**	0.651**	0.521**	0.550**	0.629**	0.563**	1.000						
A6_M2 (11)	0.691**	0.031	0.720**	0.599**	0.762**	0.772**	0.738**	0.593**	0.602**	0.594**	1.000					
A9_M2 (12)	0.620**	-0.015	0.616**	0.621**	0.698**	0.654**	0.701**	0.605**	0.591**	0.602**	0.697**	1.000				
AS3_M2 (13)	0.683**	0.137**	0.549**	0.517**	0.714**	0.567**	0.663**	0.695**	0.605**	0.531**	0.581**	0.587**	1.000			
AS6_M2 (14)	0.493**	0.030	0.362**	0.388**	0.513**	0.513**	0.489**	0.456**	0.453**	0.425**	0.458**	0.422**	0.467**	1.000		
CCIL_1_M2 (15)	0.591**	0.059*	0.583**	0.525**	0.701**	0.624**	0.732**	0.607**	0.689**	0.547**	0.735**	0.669**	0.565**	0.359**	1.000	
CCIL_4_M2 (16)	0.392**	0.024	0.330**	0.263**	0.371**	0.434**	0.432**	0.279**	0.284**	0.378**	0.534**	0.456**	0.262**	0.258**	0.407**	1.000
CCIL_6_M2 (17)	0.706**	0.110**	0.617**	0.582**	0.794**	0.648**	0.755**	0.636**	0.624**	0.497**	0.645**	0.577**	0.662**	0.428**	0.647**	0.235**
CCIL_8_M2 (18)	0.653**	0.198**	0.601**	0.506**	0.731**	0.654**	0.712**	0.600**	0.632**	0.589**	0.688**	0.625**	0.658**	0.479**	0.666**	0.476**
DRC1_M2 (19)	0.490**	0.016	0.493**	0.422**	0.620**	0.554**	0.613**	0.568**	0.585**	0.653**	0.667**	0.622**	0.493**	0.378**	0.592**	0.441**
DRC2_M2 (20)	-0.035	0.052	0.099**	0.075**	-0.009	-0.076**	0.013	0.043	0.084**	0.072*	0.095**	0.173**	0.003	-0.035	0.154**	-0.003
DRC4_M2 (21)	0.599**	-0.026	0.631**	0.579**	0.718**	0.673**	0.713**	0.569**	0.673**	0.566**	0.762**	0.708**	0.541**	0.448**	0.740**	0.415**
DRC7_M2 (22)	0.553**	-0.041	0.556**	0.506**	0.651**	0.557**	0.677**	0.598**	0.625**	0.603**	0.613**	0.604**	0.591**	0.512**	0.659**	0.303**
A4_M3 (23)	0.060*	-0.012	0.080**	-0.008	0.059*	0.044	0.146**	0.077**	0.030	-0.076**	0.092**	-0.012	0.082**	0.024	0.047	0.059*
A7_M3 (24)	0.617**	-0.008	0.566**	0.581**	0.661**	0.628**	0.651**	0.523**	0.490**	0.444**	0.595**	0.592**	0.548**	0.338**	0.546**	0.248**
AS5_M3 (25)	0.501**	-0.077**	0.511**	0.521**	0.547**	0.489**	0.493**	0.491**	0.473**	0.393**	0.437**	0.376**	0.544**	0.332**	0.424**	0.162**
CCIL_2_M3 (26)	0.575**	-0.025	0.519**	0.582**	0.628**	0.554**	0.533**	0.528**	0.486**	0.420**	0.546**	0.464**	0.484**	0.302**	0.455**	0.170**
DRC10_M3 (27)	0.482**	-0.004	0.382**	0.434**	0.530**	0.514**	0.586**	0.490**	0.496**	0.430**	0.470**	0.461**	0.427**	0.300**	0.501**	0.221**
DRC11_M3 (28)	0.610**	-0.031	0.554**	0.595**	0.666**	0.612**	0.662**	0.579**	0.500**	0.456**	0.610**	0.558**	0.526**	0.345**	0.543**	0.299**
DRC3_M3 (29)	0.475**	-0.118**	0.491**	0.561**	0.623**	0.519**	0.597**	0.524**	0.520**	0.519**	0.561**	0.553**	0.430**	0.312**	0.587**	0.305**
DRC6_M3 (30)	0.083**	-0.023	0.047	-0.005	0.097**	0.083**	0.157**	0.083**	0.138**	0.000	0.170**	0.012	0.115**	0.132**	0.131**	0.159**
AS3_M3 (31)	0.591**	-0.115**	0.530**	0.560**	0.637**	0.573**	0.621**	0.624**	0.555**	0.504**	0.609**	0.549**	0.537**	0.374**	0.611**	0.319**

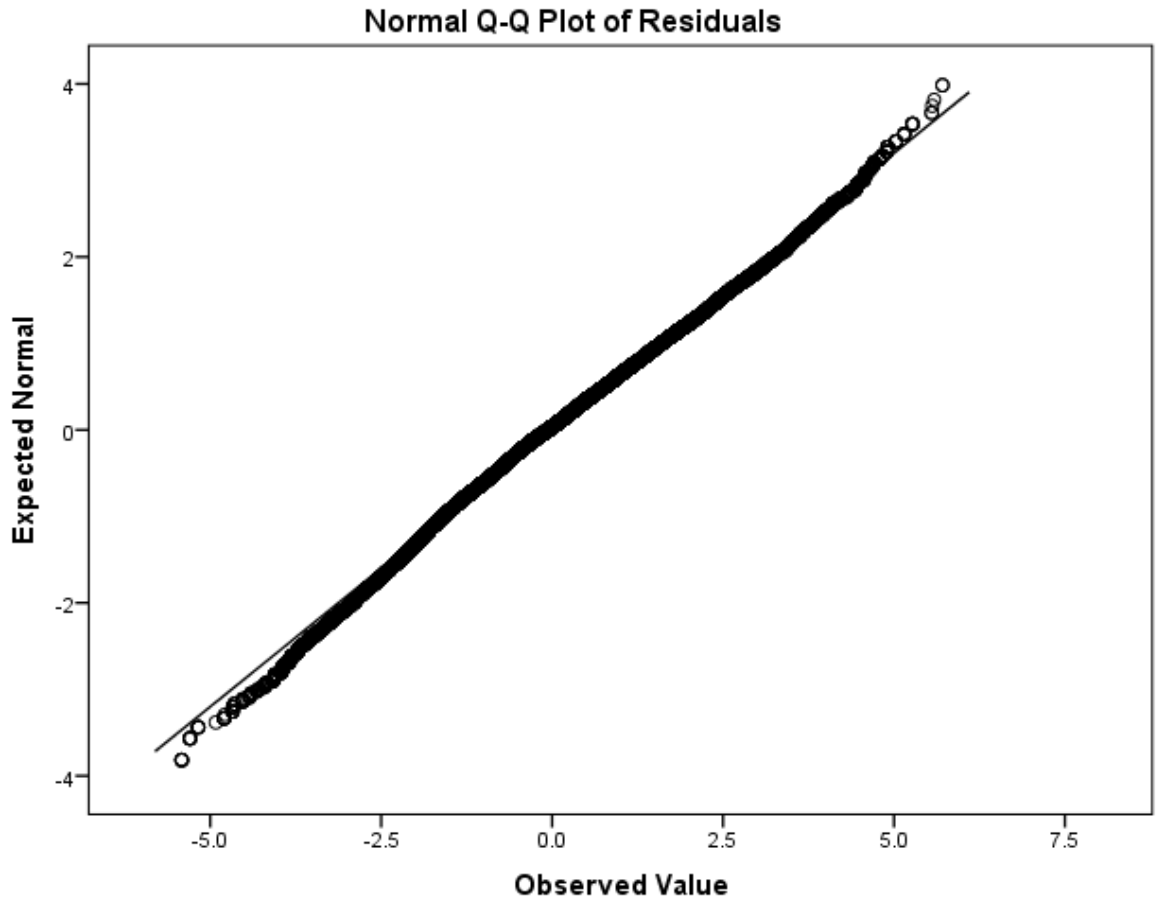
** $p < 0.01$; * $p<0.05$

Table 7. *Spearman's correlation coefficient among non-professionals (Cont'd)*

	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
A1_M1 (1)															
A5_M1 (2)															
A8_M1 (3)															
AS4_M1 (4)															
CCIL_3_M1 (5)															
CCIL_5_M1 (6)															
CCIL_7_M1 (7)															
DRC5_M1 (8)															
DRC8_M1 (9)															
DRC9_M1 (10)															
A6_M2 (11)															
A9_M2 (12)															
AS3_M2 (13)															
AS6_M2 (14)															
CCIL_1_M2 (15)															
CCIL_4_M2 (16)															
CCIL_6_M2 (17)	1.000														
CCIL_8_M2 (18)	0.606**	1.000													
DRC1_M2 (19)	0.466**	0.585**	1.000												
DRC2_M2 (20)	0.026	-0.085**	0.176**	1.000											
DRC4_M2 (21)	0.587**	0.631**	0.661**	0.116**	1.000										
DRC7_M2 (22)	0.551**	0.666**	0.549**	0.011	0.739**	1.000									
A4_M3 (23)	0.141**	0.123**	0.078**	-0.017	0.068*	-0.002	1.000								
A7_M3 (24)	0.600**	0.538**	0.396**	-0.138**	0.549**	0.488**	0.047	1.000							
AS5_M3 (25)	0.468**	0.413**	0.362**	-0.065*	0.437**	0.443**	0.014	0.667**	1.000						
CCIL_2_M3 (26)	0.520**	0.459**	0.371**	-0.068*	0.455**	0.344**	0.171**	0.776**	0.659**	1.000					
DRC10_M3 (27)	0.484**	0.506**	0.422**	-0.053	0.490**	0.484**	0.002	0.710**	0.656**	0.627**	1.000				
DRC11_M3 (28)	0.567**	0.539**	0.443**	-0.098**	0.592**	0.488**	0.090**	0.821**	0.643**	0.787**	0.732**	1.000			
DRC3_M3 (29)	0.522**	0.509**	0.438**	0.045	0.623**	0.573**	0.048	0.678**	0.557**	0.643**	0.695**	0.724**	1.000		
DRC6_M3 (30)	0.091**	0.084**	0.076**	-0.037	0.056	0.061*	0.564**	0.120**	0.155**	0.174**	0.052	0.093**	0.086**	1.000	
AS3_M3 (31)	0.546**	0.521**	0.496**	-0.003	0.614**	0.550**	0.036	0.731**	0.669**	0.714**	0.688**	0.739**	0.702**	0.098**	1.000

** $p < 0.01$; * $p<0.05$

Appendix K. *Q-Q Plot*



Appendix L. *Demographic forms for raters*

Music for Life engagement pilot study

Age	*Gender*
Ethnicity	*If completed the UK* *White British* ☐ ☐ *Black British* ☐ *Asian British* *Other Ethnic group _Chinese_____* ☐ *If completed in Hong Kong* *Chinese* *Other Ethnic group _____*
Do you play musical instrument?	☐☐ *Yes x No*
Are you in a singing group?	☐☐ *Yes x No*
Have you ever been in a singing group?	☐☐ *Yes x No*

Appendix M. *Research diary extract*

26 January 2018

I spent some time today reflecting on the process which led to me adopting this research question, and below were some of my reflections:

My research question arose from my earlier experience as part of a research team, investigating the effect music intervention have for people with dementia. Through being in that experience I was made aware of the different type of measures in measuring the "effectiveness" of the interventions, this included brief cognitive tests (MMSE, AMT), psychological outcome measures (PHQ-9, GDS, RAID…), sleep measures (PSQI) and quality of life measures (QOL-AD). Interestingly, the research I was involved in found little effect music intervention has on those outcomes. On the other hand, from my observation, as well as, the feedback that I gathered from PWD and carers, I felt that the interactions and relationship that participants gained from the inventions were overlooked. These factors were not able to be captured by the traditional outcome measures that were used.

This made me wonder whether there is an alternative way of measuring the effectiveness of music intervention or potentially other types of intervention. This sparked my curiosity, and I proposed to my supervisor about developing a tool that would measure these interactions.

My selected topic evolved as I began reading learn more about person-centred care, as well as reading about empowering PWD to make their own personal choice. This is particularly difficult for people at the advanced stages of dementia, as they are not often able to express themselves. On the other hand, understanding their choices is important in maintaining personhood.

In my attempt to identify and capture a way to measure interactions that would also consider PWD's choices, I came across the concept "engagement"; a concept of social interaction. Sabat suggested that PWD that does not participate or engage in the intervention are often misinterpreted as challenging behaviour, but in fact, it might be a result of the individual having a strong sense of self that communicates its unwillingness to engage in the interaction or activity. Sabat proposed that this sense of self should be attained with the assistant of carers and professionals. I began to wonder whether creating a tool to measure engagement could potentially benefit in staff to understand PWD that were unable to express their needs verbally.

26th November 18

I conducted five data collections this morning, in a room, I booked in the team research office. What I remembered most today, was a young "shy" looking male participants. Unlike other participants I have interacted with, this person did not maintain any eye contact with me majoring of the time. Interestingly, as I was explaining to him about the purpose of the study, he replied repeatedly by "Yup, Yup, Yup". This response was very loud and seemed very "robotic", so loud to the extent that a colleague from the office room next door came in to ask him to be quieter. Throughout the ratings, he seemed to lose his interest very early, where he often looked away from the table. This did raise my curiosity about whether he might be on the autism spectrum. As people with autism often struggle with social cues, I wondered how difficult it would be for people with autism to rate abstract concepts such as engagement. Thinking about this made me worry that this particular participant might affect my findings, but at the same time feeling guilty that I have this thought. This experience has made me realised that I have not really thought through enough about the exclusion criteria.

14th February 19

I carried out a data collection in a dementia carer group. When I went to the group, I was surprised that the majority of the carers were in their 70s and above. Four carers agreed to participate in the research. As they were not very familiar with the computer table, it took me a long time to express to them the steps and procedure of the study. I noticed that it also took two of the carers more extended period of time than other participants for them to respond and rate changes in engagement on the video. This made me wondered the type of analysis plan that would need to consider an individual's reaction time as a variable.

Furthermore, after one of the carers completed the ratings, she said to me that she found the experience of watching the videos very upsetting. She said to me that she noticed a lot of the PWD in the videos she watched looked "bored" and they were just sitting there. She stated that the facilitator in the videos should plan more exciting exercise for the PWD so that they could enjoy and have more fun. What struck me most was she then said to me that she "does not want to end up like that and not be cared for". This made me realised that observing and ratings is a subjective experience and can never be completely objective. We all bring our interpretation of reality into the video we watched or rated. Watching a "non-responsive" individual can be a difficult experience, as it can often trigger unwanted feelings (e.g. fears and anxiety), as well as existential questions (e.g. fears of death) to our conscious. Those feelings or thoughts are those that often we try very hard to repress. This also made me wonder whether ratings from the raters were, in fact, PWD's "actual" level of engagement, or the projection of the unwanted feelings that the rater projected onto the PWD when watching the video.

Appendix N. *End of study notification letter to ethics panel*

This has been removed from the electronic copy

Appendix O. *Participant findings summary letter*

[Participants]

[Street Address] [Town]
[County]

[Postcode]

[Date]

Dear [Participants],

This is a letter to say thank you for generously giving your time to take part in my research project earlier this year. Your participation helped us to explore whether looking at whether it is possible to use a tablet-computer based rating scale to rate the level of engagement for people with dementia. Your participation helped us to collect data which suggests that with further research, the tablet-computer based scale could potential be helpful people with dementia. Below is a brief summary report of what we found out.

Aim

To develop an observational rating tool with a user-friendly operational protocol to measure the level of engagement of people with advanced dementia: The Video Analysis Scale of Engagement (VASE). The study was carried out in two stages: the protocol development stage and the research stages stage.

In the development stage invited some of the volunteers to try rating a YouTube video of a person with dementia in carrying out a music activity. The volunteers were asked to provide feedback of their experience in using the tool, as well as asking them question about what engagement looks like for them. Whereas in the testing stage, we asked raters of general public and of professional backgrounds to observe and rate level of engagement in music-based activity (Music for Life and passive music listening) in residential care.

What we learnt

We found four main categories of behaviour expressions that people consider as engagement in dementia settings. These included: a) facial expressions; b) bodily movement and verbal articulations; c) attention and awareness of activity; and d) emotional responses. The findings helped us better understand

the concept of engagement and how this concept can be as a measure for advance dementia settings, especially in music-based interventions.

In addition, the research stage showed that concepts like engagement, which changes moment-by-moment can be capture using a dynamic, continuous ratings scale. There are also some support for the suggestion that the VASE is able to distinguish between the two types of music activities (passive and active). Findings suggests that with adjustment and further investigation, the tool could be used in advanced dementia care activities.

Thanks again for taking part.

Yours sincerely,

Trainee Clinical Psychologist Canterbury Christ Church University

Appendix P. *Author guideline notes for chosen journal*

This has been removed from the electronic copy