

A randomized controlled trial comparison of three music therapy formats on cognitive function and psychological well-being in normal aging

Anthony M. A. Mangiacotti, PhD^{1,2}, Emma V. Ward, PhD², Sophie Williams, MA³, Clare Barone, MA³, Khatima Ataye, MSc⁵, Maria Luisa Asue, PhD², Afsheen Behram, MSc⁵, Ming Hung Hsu, PhD⁴, Elena Lovato, MSc⁵, Alessandro Zandonà, PhD⁵, Fabia Franco, PhD²

¹ University of West London, Psychology Department, London, UK

² Middlesex University, Psychology Department, London, UK

³ Methodist Homes (MHA), Music Therapy, Derby, UK

⁴ Norwegian Academy of Music, Department of Music Education and Music Therapy, Oslo, Norway

⁵ Università degli Studi di Padova, Department of Information Engineering, Padua, Italy

Corresponding author

Anthony M. A. Mangiacotti, PhD.

E-mail: anthony.mangiacotti@uwl.ac.uk

Abstract

Objectives: Prior studies suggest a benefit of music-based activities for supporting cognitive function and psychological well-being in normal aging, but few have rigorously examined the effect of music therapy (MT) interventions in this population, or contrasted delivery format. We conducted a randomized controlled trial study investigating the effects of one-to-one, small-group, and large-group MT on cognitive functions and well-being in 62 older adults without cognitive impairment.

Methods: Participants were randomly assigned to one-to-one ($M_{age}=79.71$), small-group ($M_{age}=82.00$), or large-group ($M_{age}=79.83$) conditions, and received structured MT sessions over 20-weeks, while a passive control group ($M_{age}=74.56$) received no MT. Pre-post intervention measures assessed cognitive functions (memory, attention, executive function, verbal fluency) and well-being (general well-being, depression, sleep quality).

Results: Mixed ANOVAs indicated improvements pre-post intervention on several cognitive measures and depressive symptoms (main effects of Time) that varied as a function of MT format (Time x Group interactions). All MT groups showed improvements in cognitive functions pre-post intervention, while many cognitive scores declined in the control group. Further, all intervention groups showed a significant reduction in depressive symptoms pre-post intervention, while there was no change in the control group. Dynamic Bayesian Network analysis corroborated these trajectories and highlighted the influence of cognitive reserve and executive functions on intervention outcomes.

Discussion: This study provides evidence for the beneficial effect of MT on cognitive function and well-being in normal aging, suggesting that both individualised and group formats are advantageous. This has implications for supporting healthy aging through socially prescribed standardised protocols.

Clinical Trial Registration: NCT04851028

Keywords: cognitive decline, healthy aging, music interventions, randomized controlled trial, social prescribing

Introduction

The World Health Organization defines healthy aging as developing and maintaining the functional ability that supports well-being in older age. Identifying factors that promote healthy aging is increasingly important given the global rise in life expectancy. By 2050, the number of people aged 60+ is projected to double from 1 billion in 2020 to nearly 2.1 billion, while those aged 80+ are expected to triple to about 426 million (World Health Organization, 2024). While dementia cases are rising, a significant proportion of older adults do not have a diagnosed neurocognitive disorder. Normal aging involves gradual cognitive decline (Panico et al., 2023), and reductions in emotional well-being (e.g., Chen et al., 2021; Li et al., 2025), so supporting cognition and well-being in this population is a critical health priority (Sun et al., 2023).

A major modifiable factor in cognitive aging is cognitive reserve, developed over the lifespan through education, occupation, and stimulating leisure activities (Panico et al., 2023). Social engagement, physical activity, and cognitive stimulation also protect against decline, making them key targets for prevention (Oosterhuis et al., 2023). Indeed, social prescribing, an approach that connects people to community resources and activities to support social, emotional, and practical needs, is increasingly promoted by policymakers to improve the lives of older individuals. Common activities include arts and crafts, physical exercise, and music-based programs (Costa et al., 2021). Emerging evidence suggests that such initiatives are a valuable preventative measure, linked to lower use of public health services (National Academy for Social Prescribing, 2024).

Interest in music therapy (MT) has grown for its potential to promote cognitive, emotional, and psychosocial well-being in older adults (Jang & Kunde, 2021; Mangiacotti et al., 2024a; 2024b). MT is a psychological intervention involving the professional use of music and its elements to enhance quality of life, physical health, and emotional health (WFMT,

2011). A qualified therapist facilitates activities such as playing instruments, singing, or listening. Multiple approaches exist (Darrow, 2008), but improvisation, spontaneous music-making between therapist and client, fosters active participation and has been shown to enhance social skills, intrinsic motivation, cooperative learning, concentration, creativity, self-awareness, and self-regulation (Biasutti, 2015). Affect matching is another key element, as cognitive gains are linked to music that mirrors the listener's emotional state (Franco et al., 2014; Ward et al., 2021). In improvisational MT, this means the therapist creates music that aligns with the recipient's inner state (Ansdell & Pavlicevic, 2005).

MT has been associated with benefits in older adults with and without neurocognitive decline, including improvements in cognitive functions (e.g., attention, working memory), social functioning (e.g., communication, cooperation, empathy), mood (e.g., reduced depressive symptoms and anxiety), quality of life, and reductions in behavioural symptoms (Hsu et al., 2015; Jang & Kunde, 2021; Ma & Ma, 2023; Wang et al., 2023b). Neuroimaging studies further show engagement of neural networks involved in executive control, memory, and sensory monitoring, shedding light on the mechanisms underlying these effects (Faber et al., 2025).

Most systematic reviews and meta-analyses on MT have focused on populations with mild cognitive impairment or dementia. Bleibel et al. (2023) reported that active interventions yield greater cognitive benefits than passive listening, but noted inconsistencies due to methodological variability (e.g., session frequency, outcome measures). Similarly, de Witte et al. (2022) stressed the need for more robust, standardised designs (see also Wang et al., 2023a; Wang et al., 2023b). Addressing these limitations, Mangiacotti et al. (2024a) conducted a randomized controlled trial (RCT) with 42 older adults ($M=86.25$ years) with mild to moderate cognitive impairment. Participants were randomized to a structured sixteen-week one-to-one MT or storytelling protocol, with cognitive, physiological, and biomarker assessments. MT led

to significant improvements in cognitive function (executive functions, attention, visuospatial abilities) and well-being (reduced depressive symptoms, improved quality of life) compared to storytelling. Further, reductions in the cortisol/DHEA ratio and increases in respiratory sinus arrhythmia suggested improved stress modulation and parasympathetic activity in the MT group. Although focused on a clinical population, the study provided a rigorous framework for evaluating MT effects more broadly.

Few studies have focused on MT in normal aging, but a recent review by Zhou et al. (2025) suggests that music-based activities can yield positive effects on cognitive function in this population compared to other activities (e.g., reading, painting). Specifically, they reported small to moderate improvements in global cognition and executive functions as a result of MT, but noted limitations such as design variability, small samples, heterogeneous protocols, and varied outcome measures (see also Xu et al., 2017). Importantly, preventive approaches in normal aging extend beyond cognition. Well-being is also a relevant target, as age-related changes in social connectedness, lifestyle, and mood can affect quality of life (Ma & Ma, 2023; Jang & Kunde, 2021), and are directly linked to cognitive decline (Chen et al., 2021; Li et al., 2025; Wang et al., 2023a). Studies have shown the benefits of community music interventions on well-being makers, such as a reduction in depressive symptoms, anxiety, and improved quality of life, possibly mediated by strengthened social relations, enhanced self-confidence and self-esteem, and reduced social isolation (Ma & Ma, 2023). These pathways may support cognitive functioning, as socially and cognitively enriching musical activities involve complex, multimodal demands (e.g., attention, memory, executive control, social coordination) (Biasutti & Mangiacotti, 2018; 2021; James et al., 2024). Jang and Kunde (2021) and Raja et al. (2025) also report preliminary benefits on emotional well-being, physical fitness, and social functioning in healthy older adults.

Sleep quality has also been examined in aging, given its association with quality of life and cognitive functioning (Chen et al., 2021). Recent systematic review and meta-analytic evidence suggest that music interventions may improve sleep in older adults (Li et al., 2025), but the quality of evidence is low due to methodological limitations, including varied study designs, intervention characteristics, and outcome measures. Additionally, participant cognitive status has not been consistently reported, with some studies mixing healthy and clinical samples, limiting conclusions around sleep-related benefits in cognitively healthy older adults (Chen et al., 2021; Li et al., 2025).

Overall, MT shows potential to support cognitive and emotional health in normal aging, but there is a need for more robust, standardised investigations. Further, there is a lack of studies directly comparing different MT delivery formats, such as contrasting individual versus group-based activities. This is crucial to understand the mechanisms underlying any benefits, specifically, the extent to which improvements are a direct function of MT, and the contribution of social-context effects. Understanding the effectiveness of different formats is also essential to optimise standardised protocols. This study aimed to address these gaps.

We report a structured RCT study to examine the effect of three MT formats on cognitive function and psychological well-being (general well-being, depression, sleep quality) in healthy older adults. Reflecting formats in which MT is typically applied in clinical and community settings, we compared one-to-one, small-group, and large-group MT against a passive control group. While the MT protocol was constant, the delivery format systematically varied the social context (individual versus group-based interactions), and the degree of therapist individualisation and patient-centred focus. Our objectives were to evaluate whether MT can meaningfully support cognitive functions and well-being in normal aging and compare the effectiveness of the different formats.

Methods

Participants

An a-priori power analysis was conducted using G*Power version 3.1.9.7 (Erdfelder et al., 1996). The calculation was based on the effect size ($f(U)=0.497$) observed on the primary outcome measure in Mangiacotti et al. (2024a), the Music Cognitive Test. Alpha was set at .05, power ($1-\beta$) at 0.80, with four groups and two repeated measures (pre-post design). This yielded a required sample size of 52 participants. To allow for attrition, we overrecruited to ensure that we would have at least the required number of participants. Our final sample size comprised 62 participants who completed the study in full (19.2% over-recruiting).

Participants were recruited from five Methodist Homes (MHA) community centres within and outside Greater London to ensure a broadly representative sample. MHA is a UK-wide charity providing services, including MT, to older adults. Additional participants were recruited from the University of the Third Age (U3A). Ethical approval was granted by the Middlesex University Research Ethics Committee. The study, including pre- and post-intervention data collection, ran from November 2022 to May 2023.

Inclusion criteria were: i) age ≥ 65 years; ii) no cognitive impairment (Mini Mental State Examination, MMSE, Folstein et al. 1983, score ≥ 25); iii) fluency in English; iv) no significant hearing impairment that could affect participation in MT activities. Exclusion criteria were: i) severe motor deficits limiting engagement in activities; ii) participation in a cognitive training or MT programme within the last six months; iii) received a cognitive assessment within three months of the pre-intervention test battery. Initially, 413 individuals received information about the study through email and/or meetings at community centres, and 148 individuals expressed interest. Of these, 62 individuals did not meet the eligibility criteria or did not later respond (Figure 1). Thus, 86 participants were initially enrolled in the study and completed the pre-intervention test battery prior to randomisation to i) one-to-one, ii)

small-group (n=4-8), iii) large-group MT ($n \geq 15$)¹, and iv) passive control conditions. Randomisation was done using a computerised procedure by a researcher blind to group allocation. All groups were comparable at baseline on key demographic and cognitive variables. Of the 86 enrolled, 62 participants completed the study in full, including the post-intervention test battery, and were included in the analysis. In total, 24 participants (28%) dropped out or were excluded: two withdrew due to medical reasons, seven withdrew without giving a reason, three commenced new therapies and became ineligible, and 12 had insufficient attendance (<50% of MT sessions) (Figure 1). Table 1 summarises demographics and baseline scores in the final sample; there were no significant differences between groups in demographics, general health, cognitive reserve, or on the primary cognitive or well-being assessments.

Baseline scores were consistent with a cognitively healthy sample and, across most outcomes, indicated scope for improvement to be measured. The one-to-one group showed slightly better baseline performance across cognitive outcomes compared to other groups, but not significantly different. Specifically, this group had the highest baseline MCT score ($M=48.82/52$), so there was less room for measurable change over time on this task in this group.

Design and Procedure

The intervention comprised 20 weekly MT sessions delivered once a week over approximately five months, with all groups running in parallel. Measures for both intervention and control participants were collected up to two weeks before the first session and within one week of the final session.

Intervention phase:

MT sessions were delivered by five HCPC-registered music therapists, each with at least two years' experience working with older adults. All received in-depth training on the study and followed the same protocol. Participants remained with the same therapist throughout, and sessions were held in comfortable, well-lit, acoustically isolated community centre rooms, free from distractions.

Here we provide a summary of the MT protocol, and full details are in Supplementary Material, including a protocol checklist based on Robb et al.'s (2025) reporting guidelines for music-based interventions. Each session (~40 min) began and ended with specific 'hello' and 'goodbye' songs, and therapists interactively engaged participants through singing, playing instruments, and asking questions to prompt reflection on the music and listener experiences. Music was chosen based on the participant's history or requests, from a database curated by the MT team. All therapists had access to the same number and quality of musical instruments and technologies. In line with established MT frameworks (Hsu et al., 2015), we adopted an improvisational, affect-matching approach (see Introduction). Therapists responded musically to participants' behaviours while implicitly attending to their affective state (e.g., alertness, apathy, mood). This involved adapting musical rhythm, melody, and intensity to participants' actions and affective cues, facilitating active exploration of sound and interaction (Hsu et al., 2015; Santi & Zorzi, 2016). In other words, there was continuous therapist modulation of musical parameters (melody/pitch, rhythm, tempo, timbre, dynamics, and harmony) in direct response to participants' vocal, instrumental, and bodily actions, forming a structured musical dialogue. Periods of silence were used to facilitate listening, turn-taking, and regulation. Verbal exchange, intermittent eye contact, and guided physical engagement (e.g., instrument-related movement, body movement, or dance-like actions) were incorporated to support embodied participation.

In one-to-one sessions, tasks were carried out directly between the participant and therapist. Each participant received the therapist's full attention, enabling focused exploration of feelings and experiences based on individual preferences. In small-group sessions, the therapist facilitated music interaction tailored to the group while still engaging with each participant individually. Large-group sessions included activities such as group singing and instrument circles, with the therapist addressing the whole group. Both small- and large-group formats encouraged social interaction, though the small-group setting provided a more personal atmosphere and greater opportunities for interpersonal contact.

Baseline and post-intervention tests:

Before the intervention, background information was collected, and participants were screened for cognitive impairment. Baseline measures were obtained in a separate session using a standardised test battery (see Materials), assessing cognitive functions and psychological well-being (~45 min). Within one week of the final MT session, participants completed the test battery again. All assessments were conducted individually and privately, administered by five researchers experienced in neuropsychological testing of older adults, and who were blind to the conditions of the participants. All researchers received training on the specific test battery.

Materials

Background and screening tests

The MMSE (Folstein et al., 1983) was used to screen for cognitive impairment. This widely used standardised test assesses general cognitive function, with scores ranging from 0 to 30. Scores <24 indicate cognitive impairment. Only participants scoring 25 or above were included in the study. The MMSE shows good internal consistency ($\alpha=.60-.85$).

The Cumulative Illness Rating Scale (CIRS; Linn et al., 1968) assesses physical illness severity across 14 categories, based on medical history and symptoms. It provides a total score

(0–56), a Comorbidity Index (CIRS_C: number of categories with score ≥ 2), and a Severity Index (CIRS_S: overall burden). Higher scores indicate greater severity. Reliability scores range from .55 to .91.

The Cognitive Reserve Index Questionnaire (CRIq; Nucci et al., 2012) was used to assess cognitive reserve via a structured interview. It combines data on education, occupation, and leisure into a single index ($M=100$, $SD=15$), and has a reliability score of $\alpha=.62$.

Pre-post intervention outcome measures

Cognitive tests

The Music Cognitive Test (MCT; Mangiacotti et al., 2022) was used as the primary cognitive outcome measure for its sensitivity to the specific cognitive skills targeted by music engagement. It correlates with other widely used cognitive assessment tools such as MMSE and MoCA (Nasreddine et al., 2005), and displays excellent sensitivity in identifying cognitive impairment according to MMSE and MoCA diagnostic criteria (99.4% and 93.0%, respectively). Scores >45 (range: 0–52) indicate cognitive normality. The MCT has good internal consistency ($\alpha=.90$) and test–retest reliability (Pearson’s $r=.94$, over one month).

Verbal Fluency (VF) was assessed through the MCT subscale for assessment of lexical ability, verbal fluency, and strategic organisation. It is a modified version of the Verbal Fluency Test (Machado et al., 2009), in which participants generate as many words as possible starting with a designated letter in 60 seconds. The score reflects the number of correct, non-repeated words produced. This test shows good internal consistency ($\alpha=.78$) and test-retest reliability ($ICC=.89$, one month).

The Tangled Figure Test (TFT; Mondini et al., 2011) was used to assess spatial cognition, and executive and naming difficulties. Participants have 4 minutes to identify as many items as possible from a tangled picture containing 50 items. The total score is the number of correctly identified items. This scale has good test–retest reliability (Pearson’s

$r=.89$; one month), and allows error categorisation into Visual (TFT-V) and Semantic (TFT-S) errors, Anomia (TFT-A), and Repetitions (TFT-R).

The Trail Making Test (TMT-A; Mondini et al., 2011) was used to assess sustained and selective attention. Participants are instructed to connect numbers in ascending order across a sheet. The score is based on completion time in seconds. This scale presents good internal consistency ($\alpha=.98$) and test–retest reliability (Pearson’s $r=.96$; one month).

Mood and Behavioural Tests

The Geriatric Depression Scale (GDS; Burke et al., 1991): The brief (15-item) version was used to assess depressive symptoms. Scores can range from 0-15, with scores >5 indicating possible depressed mood. The GDS demonstrates excellent internal consistency ($\alpha=.81-.92$) and test-retest reliability over two weeks ($ICC=.83$).

The Well-being of Older People Scale (WOOP; Hackert et al., 2020) was used as a general well-being index. Participants rate items within a list of nine well-being factors (physical health, mental health, social contacts, support, acceptance and resilience, feeling useful, independence, making ends meet, and living situation) on how important they are to their overall well-being. Higher scores indicate greater self-perceived well-being. The scale demonstrates high test-retest reliability ($ICC=.89$; one week).

The Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989) was used to evaluate sleep quality. This self-report questionnaire includes seven subcategories: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction. The scale has good internal consistency ($\alpha=.83$) and good test–retest reliability (Pearson’s $r=.85$, one month).

Analysis

Analyses were performed using IBM SPSS 25 and R 4.3.2. An alpha level of .05 was used, and we report partial eta squared (η_p^2) effect sizes for ANOVA effects. All data were tested for normality (Kolmogorov–Smirnov test), homogeneity of variance, and sphericity. Scores on each measure were first analysed using a 2(Time) x 4(Group) mixed ANOVA, with Time (pre-post intervention) as the within-subject factor and Group (one-to-one, small-group, large-group, control) as the between-subjects factor. We then conducted a-priori planned comparisons pre-post intervention for all measures. Given (nonsignificant) differences in age between groups, we also performed the analysis including age as a covariate (provided in Supplementary Material). Finally, as a further exploratory analysis, we conducted a Dynamic Bayesian Network (DBN) analysis to unravel the probabilistic dependencies among variables over time. Two types of DBN were estimated: a bootstrapped network based on 50-random samples with replacement, which identified dependencies recurring more consistently across resamples, and a single-fit network obtained from the full dataset without resampling, for which conditional probability tables (CPTs) were examined. We report a general overview of the outcomes of the DBN analysis and full details in Supplementary Material. The de-identified raw data and analysis code are available at <https://doi.org/10.22023/mdx.29979994>

Results

See Table 1 for sample demographics and baseline scores, and Table 3 for pre-and post-intervention means.

Preliminary analyses

The outcomes of the 2(Time) x 4(Group) mixed ANOVAs are given in Table 2 (see Table 3 for planned comparisons pre-post intervention). Significant main effects of Time indicated improvement pre-post intervention on measures of attention, visuospatial abilities,

executive functions, and verbal fluency (TMT-A, TFT, VF, TFT-S, marginal effect on the MCT), and a reduction in depressive symptoms (GDS). Time x Group interactions were observed for several cognitive measures (MCT, TMT-A, TFT, and TFT-V), indicating that changes on these tests varied across MT format (examined further below).

Effects on cognitive functions

Figure 2 Panel A depicts pre-post intervention changes on cognitive measures in each group. The one-to-one group showed a reliable improvement in attention (TMT-A, $p=.017$) and visuospatial abilities (TFT, $p<.001$), with a significant reduction in the number of TFT semantic errors (TFT-S, $\Delta = -0.53$, 95 % CI $[-0.94, -0.12]$, $p=.012$, $\eta_p^2=.105$). The small-group condition showed significant improvement in general cognitive function (MCT, $p<.001$), attention (TMT-A, $p=.024$), and verbal fluency (VF, $p=.048$), and the large-group showed significant improvement in general cognitive function (MCT, $p=.034$) and visuospatial abilities (TFT, $p<.001$). In all other cases, measures were stable in the intervention groups over time. However, the control group showed a reliable decline in general cognitive function (MCT, $p=.014$), and on aspects of the TFT. Specifically, examination of error types on this test revealed a marked increase in visual errors ($\Delta = +2.69$, 95 % CI $[0.85, 4.53]$, $p=.005$, $\eta_p^2=.128$) and anomia errors ($\Delta = +0.56$, 95% CI $[0.14, 0.99]$, $p=.010$, $\eta_p^2=.108$) in the control group.

Effects on psychological well-being

Figure 2 Panel B shows pre-post intervention changes in well-being measures by group (Table 3 for planned pre-post intervention comparisons). All intervention groups showed a significant reduction in depressive symptoms (GDS, all $p<.05$), while the control group showed no change. Well-being (WOOP) and sleep quality (PSQI) indices remained stable over time in all groups.

Dynamic Bayesian Network

DBNs are graphically represented as Directed Acyclic Graphs whose nodes are the variables measured pre-post intervention; ‘_P’ representing post intervention, and whose edges are their probabilistic dependences. Across both networks, Group emerged as an upstream driver, most consistently linked to cognitive reserve (CRIq), and in the bootstrap also to global cognition (MMSE) and cumulative illness (CIRS); CPTs supported these associations (see Supplementary Material for Figures and conditional probabilities between variables). From cognitive reserve, effects extended to verbal fluency (VF) and the Tangle Figure Test (TFT), which acted as a central hub predicting attention/processing speed (TMT-A) and several post-intervention measures, including TFT and post-intervention sleep quality (PSQI) and general cognitive functions measured by MCT; MMSE also predicted post-MCT results, linking global cognition to later cognitive performance. On the well-being side, a pathway connecting CIRS to WOOP and to GDS, and a pathway from WOOP to PSQI, were observed (see Supplementary Table 8 and Supplementary Figure 2), indicating that baseline well-being predicted post-intervention well-being. Pre-post predictive links were observed across domains (TMT_A, TFT, WOOP, GDS, and PSQI), indicating that baseline levels influenced follow-up scores. Demographic effects were selective: A connection between Age and TFT appeared in the single-fit model, consistent with age-related differences in spatial cognition, while sex-related effects on sleep quality were observed only in the bootstrap. In general, the group had both a direct and indirect influence on post-intervention cognitive and well-being measures.

Discussion

This study examined the effects of three music therapy (MT) intervention formats on cognitive functions and psychological well-being in normal aging, comparing one-to-one,

small-group, and large-group MT with a passive control. All MT groups improved on various cognitive measures pre-post-intervention, while general cognitive function (Music Cognitive Test; MCT) declined in the control group. Specifically, the one-to-one group improved in attention (Trail Making Test; TMT-A) and visuospatial abilities (Tangled Figure Test; TFT), the small group in general cognition (MCT), attention (TMT-A), and verbal fluency (VF), and the large group in general cognition (MCT) and visuospatial abilities (TFT). Other cognitive measures remained stable over time in intervention groups. All intervention groups showed reduced depressive symptoms (Geriatric Depression Scale; GDS) pre-post-intervention, with no change in the control group. No changes were observed in sleep quality (Pittsburgh Sleep Quality Index; PSQI) or general well-being (Well-being of Older People Scale; WOOP).

Although all MT formats were beneficial, the partially differentiated outcomes suggest that the delivery format affected the nature of the improvements. The Dynamic Bayesian Network (DBN) analysis clarifies how group allocation, baseline characteristics, and outcome domains interrelate over time. MT format influenced post-intervention cognition primarily through indirect pathways. Group affected post-intervention MCT and TMT-A via baseline cognitive reserve (Cognitive Reserve Index Questionnaire; CRIq) and global cognition (Mini Mental State Exam; MMSE), which in turn fed into verbal fluency (VF) and visuospatial abilities (TFT). The TFT, assessing spatial cognition, executive, and naming abilities, emerged as a central hub linking these baseline capacities to later gains in attention and processing speed. This aligns with the ANOVA results, where the one-to-one format was associated with improvements in TMT-A and TFT, the small-group format in MCT, TMT-A, and VF, and the large-group format in MCT and TFT, while the control group declined in MCT and increased in specific TFT error types (anomia and visual errors). On the psychosocial side, a comorbidity-well-being-mood (CIRS-WOOP-GDS) pathway (see Supplementary Table 8) explains why all intervention groups, but not the control, showed reductions in depressive symptoms despite no

significant mean change in well-being or sleep quality. Well-being mediated the relationship between physical health and affective outcomes. The DBN also revealed strong pre-post predictive links in most domains (e.g., TMT-A, TFT, GDS, PSQI), indicating that baseline status shaped the magnitude of change. Overall, these patterns suggest that cognitive and well-being benefits in the MT groups were supported by specific baseline resources, particularly cognitive reserve, global cognition, and psychosocial well-being, which channelled the effects of the different formats into distinct outcome profiles.

The small and large-groups showed significant improvement pre-post intervention on the MCT, the primary measure of general cognitive function. MCT scores were stable in the one-to-one group, but it should be cautioned that this group had the highest baseline score ($M=48.82/52$), meaning less room for measurable gains. We cannot conclude that there was no improvement in general cognition in the one-to-one group. Indeed, individuals with higher baseline abilities commonly show smaller or less measurable gains from interventions (Ma & Ma, 2023; Mangiacotti et al., 2024a; Panico et al., 2023; Zhou et al., 2025). DBN findings support this interpretation, as baseline MMSE scores were directly linked to post-MCT outcomes, indicating that general cognitive functioning constrained or enabled MCT change. The one-to-one group did, however, show significant improvements in visuospatial processing and executive functioning. Interestingly, only this group showed a reduction in semantic errors on the visuospatial task (TFT). This may reflect deeper engagement and a more sustained focus in individualised sessions. For example, immediate, individualised feedback may support sustained monitoring and semantic verification processes, which could contribute to reduced semantic errors relative to group formats. The small-group condition was most effective for enhancing general cognitive function, and was the only group to show improvement in verbal fluency. This may be due to the combination of relatively individualised MT alongside meaningful social interaction (Biasutti, 2015; Oosterhuis et al., 2023). This format offered a

balance of group-based engagement with opportunities for individual participation (e.g., turn-taking and brief verbal exchanges), which may have contributed to the observed language-based outcomes (verbal fluency) alongside attention and executive improvements (Pearce et al., 2016; Tarrant et al., 2021). Large-group MT showed the largest effect size on higher-order executive functions, suggesting that a socially enriched musical environment promotes the greatest cognitive gains (Ma & Ma, 2023; Zhou et al., 2025). This aligns with studies showing that musical engagement in socially interactive settings fosters cognitive reserve, supports neural plasticity, and improves attention and memory (James et al., 2024).

When interpreting these cognitive effects, it is important to consider the potential contribution of practice effects on the measures. The test-retest reliability norms were high for all measures (see Materials), and the 5-month interval between pre- and post-intervention testing was substantially longer than that used in many of the test-retest norms (one month). It is therefore unlikely that the observed effects reflect improvements in the measures as a result of practice. Further, if this were the case, similar gains would also be expected in the control group.

The reduction in depressive symptoms pre-post-intervention in all MT groups is consistent with prior meta-analytic evidence that MT can reduce depressive symptoms in older adults (Wang et al., 2023b). The large-group condition showed the largest effect, suggesting that a socially enriched musical environment is beneficial (Ma & Ma, 2023; Zhou et al., 2025). While this condition may amplify socialisation and belonging via collective participation and interpersonal synchrony (Pearce et al., 2016; Tarrant et al., 2021), one-to-one MT may place a greater emphasis on individual emotional processing (Jang & Kunde, 2021; Raja et al., 2025). As noted by Coppi (2024) and Ito et al. (2022), large-group MT is particularly valuable in community settings, offering broad engagement that addresses both individual and socio-cultural needs.

No significant changes were found in general well-being (WOOP) or sleep quality (PSQI). However, the DBN indicated that WOOP mediated the mood pathway, with baseline well-being influencing post-intervention mood (GDS) through its link to baseline GDS. PSQI showed a baseline-to-post relationship and a connection with baseline WOOP, suggesting an influence of well-being in sleep patterns. The model also identified a direct pathway from pre-intervention executive functions (VF, TFT) to post-intervention PSQI, suggesting that stronger baseline executive control may predict better sleep outcomes. Executive functions aid in perceiving, monitoring, and regulating one's internal state (Choudhury & Halder, 2022), which is important for sleep. Broadly, while recent meta-analytic evidence suggests that music interventions may improve sleep quality in older adults (Li et al., 2025), there are a number of issues in prior studies (see Introduction), and coupled with the present null findings on the PSQI, further research is needed.

Overall, this study provides new evidence of the benefit of MT on cognitive function and well-being in normal aging across both individual and group-based formats. Using a rigorous RCT design, it adds to the growing evidence supporting MT as a valuable tool for promoting cognitive and emotional well-being in normal aging (Jang & Kunde, 2021; Raja et al., 2025; Wu et al., 2023; Zhou et al., 2025). However, caution should be applied given the lack of an active control group. Given our aim to compare MT delivery formats, we used a single passive control group as a comparator for practical reasons; it was not feasible to include three active control groups matched to the respective MT delivery format. This choice was also informed by prior studies showing benefits of MT/music interventions relative to active control groups (e.g., Baker et al., 2023; Biasutti & Mangiacotti, 2018; Feng et al., 2020; Mangiacotti et al., 2024a; Narme et al., 2014). Specifically, there is evidence that MT is advantageous for cognitive function and well-being in older adults with mild to moderate decline compared to a matched storytelling control group (Mangiacotti, 2024a), and evidence for a benefit in healthy

older adults above activities such as reading and painting (reviewed in Zhou et al., 2025). However, it is still important for future studies to include well-matched active comparators to disentangle music-specific effects from non-specific engagement effects (Ma & Ma, 2023; Zhang et al., 2024). Although we can only speculate on the mechanisms in the current study, there is evidence that music stimulates neural networks involved in attention, working memory, cognitive flexibility, spatial reasoning, and visual–motor coordination (Beaty, 2015; Biasutti, 2015), and this aligns with the cognitive benefits observed. This is also supported by evidence that the parietal cortex and prefrontal areas, critical for a range of cognitive functions, are engaged during musical activities (Beaty, 2015; Faber et al., 2025; Malekmohammadi & Cheng, 2025).

A few other limitations should be noted. First, dropout was fairly high, particularly in the large-group condition. This may reflect the characteristics of cognitively healthy older adults, who often lead active lives with diverse social interests. Unlike clinical populations, they may have less availability for sustained engagement in structured activities. Reduced opportunities for personalised interaction and therapeutic anchoring in the large-group format may have also lowered motivation for some participants in this condition, especially those with a desire for closer interpersonal contact. In addition, engagement in group-based music activities can be influenced by social dynamics (e.g., perceived cohesion, shared identity, and comfort participating in a larger group), which may differentially affect retention across delivery formats (Pearce et al., 2016; Tarrant et al., 2021). Second, despite standardisation of protocol through therapist training, individual differences in therapeutic style, song choice, and interaction approach may have influenced outcomes. This underscores the importance of detailed intervention reporting and fidelity considerations, as emphasised in recent reporting guidance for music-based interventions (Robb et al., 2025). Finally, the study did not assess participant engagement, motivation, or therapeutic alliance, factors that may explain why

certain MT formats work better for some individuals (Pearce et al., 2016; Tarrant et al., 2021). Future studies should incorporate measures related to these factors. There is also a need for longer-term follow-ups and studies to better understand mechanisms of change. It will be particularly important to directly compare the benefits of MT to other types of music activities (e.g., choirs).

To conclude, all MT formats reduced depressive symptoms and produced cognitive benefits, each with a distinct profile. One-to-one MT improved visuospatial and attentional domains, suggesting utility for individuals with early cognitive vulnerabilities, sensory challenges, or a preference for personalised formats. Small-group MT enhanced general cognition, attention, and verbal fluency, making it well-suited for those seeking cognitive stimulation in socially supportive, semi-structured settings. Large-group MT was most effective for general and executive functions, aligning with contexts where shared engagement is central, such as community centres. However, this format also had the highest dropout rate, possibly due to reduced therapeutic anchoring and diverse participant commitments. This highlights the need to tailor MT delivery to individual needs. This study responds to calls for prevention-focused trials and highlights the potential of community-based MT as a suitable approach for social prescribing. Although we did not employ a social prescribing referral pathway, the structured and scalable protocol is compatible with such schemes. National evaluations point to the utility of community social prescribing. It is therefore important to explore the integration of MT into such preventative frameworks for older adults and evaluate health-economic outcomes, value, and long-term real-world sustainability.

Funding

This work was funded by the Vivensa Foundation (RPGF2006\241).

Conflict of Interests

The authors have no conflicts of interest to declare.

Data Availability

For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission. The de-identified raw data is openly available at Figshare <https://doi.org/10.22023/mdx.29979994>.

This study was registered with <https://ClinicalTrials.gov> (Identifier NCT04851028).

Acknowledgements

We would like to express our appreciation to Methodist Homes (MHA) for their partnership and collaboration, especially the music therapists (Maria Michael, Jack Steel, Lynfa Sioufi, Jo Sarkany, Hannah Merchant), managers, and staff for their support throughout the intervention and for facilitating access to the MHA communities. We are also grateful to the University of the Third Age for their assistance in participant outreach. We thank Eka Zuniga, Ana Ivanova Zenziva, Vanessa Simoes, Jennifer Lopez, Lai Yiu Yeung, and Jwaad Hussain for their support with data collection, and to all the project's participants. Finally, we wish to honour the memory of our Principal Investigator, Prof Fabia Franco, whose vision, passion, and leadership were the driving force behind this project. Her dedication to bridging science and societal impact continues to inspire us, and this work stands as a tribute to her legacy.

References

- Ansdell, G., & Pavlicevic, M. (2005). Musical companionship, musical community. Music therapy and the process and values of musical communication. In D. Miell, R. MacDonald & D. Hargreaves (Eds.), *Musical Communication* (pp. 193). Oxford: Oxfordshire. <https://doi.org/10.1093/acprof:oso/9780198529361.003.0009>
- Beaty R.E. (2015). The neuroscience of musical improvisation. *Neuroscience & Behavioral Reviews*, 51, 108-117. <https://doi.org/10.1016/j.neubiorev.2015.01.004>
- Bleibel, M., El Cheikh, A., Sadier, N. S., & Abou-Abbas, L. (2023). The effect of music therapy on cognitive functions in patients with Alzheimer's disease: a systematic review of randomized controlled trials. *Alzheimer's research & therapy*, 15(1), 65. <https://doi.org/10.1186/s13195-023-01214-9>
- Biasutti, M. (2015). Pedagogical applications of cognitive research on musical improvisation. *Frontiers in Psychology*, 6, 614. <https://doi.org/10.3389/fpsyg.2015.00614>
- Biasutti, M., & Mangiacotti, A. (2018). Assessing a cognitive music training for older participants: a randomised controlled trial. *International Journal of Geriatric Psychiatry*, 33(2), 271-278. <https://doi.org/10.1002/gps.4721>
- Biasutti, M., & Mangiacotti, A. (2021). Music training improves depressed mood symptoms in elderly people: a randomized controlled trial. *The International Journal of Aging and Human Development*, 92(1), 115-133. <https://doi.org/10.1177/0091415019893988>
- Burke W.J., Roccaforte, W.H., Wengel, S.P. (1991). The short form of the Geriatric Depression Scale: a comparison with the 30-item form. *Journal of Geriatric Psychiatry and Neurology*, 4(3), 173-178. <https://doi.org/10.1177/089198879100400310>

- Buyse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Research*, 28(2), 193-213. [https://doi.org/10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)
- Chen, C.T., Tung, H.H., Fang, C.J., Wang, J.L., Ko, N.-Y., Chang, Y.J., & Chen, Y.C. (2021). Effect of music therapy on improving sleep quality in older adults: A systematic review and meta-analysis. *Journal of the American Geriatrics Society*, 69(7), 1925–1932. <https://doi.org/10.1111/jgs.17149>
- Choudhury, A. R., & Halder, S. (2022). The relationship between executive functions and interoceptive awareness in the elderly with perceived feelings of loneliness. *Taiwanese Journal of Psychiatry*, 36(3), 131-137. https://doi.org/10.4103/TPSY.TPSY_28_22
- Coppi, A. (2024). Developing community music in Italy: The journey. *International Journal of Community Music*, 17(1), 7-18. https://doi.org/10.1386/ijcm_00093_1
- Costa, A., Lopes, J., Sousa, C. J., Santos, O., Virgolino, A., Nogueira, P., ... & Alarcão, V. (2021). Developing a social-prescribing local system in a European Mediterranean country: a feasibility study to promote active and healthy aging. *BMC Health Services Research*, 21, 1-11. <https://doi.org/10.1186/s12913-021-07186-6>
- Darrow, A. A. (2008). *Introduction to approaches in music therapy*. American Music Therapy Association. 8455 Colesville Road Suite 1000, Silver Spring, MD 20910.
- de Witte, M., Pinho, A.D.S., Stams, G.J., Moonen, X., Bos, A.E., & van Hooren, S. (2022). Music therapy for stress reduction: a systematic review and meta-analysis. *Health Psychology Review*, 16(1), 134-159. <https://doi.org/10.1080/17437199.2020.1846580>

- Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, 28(1), 1–11.
<https://doi.org/10.3758/BF03203630>
- Faber, S., Brown, T., Carpentier, S., & McIntosh, A. R. (2025). Mapping multi-modal dynamic network activity during naturalistic music listening. *Imaging Neuroscience*, 3, 00413. https://doi.org/10.1162/imag_a_00413
- Franco, F., Swaine, J. S., Israni, S., Zaborowska, K. A., Kaloko, F., Kesavarajan, I., & Majek, J. A. (2014). Affect-matching music improves cognitive performance in adults and young children for both positive and negative emotions. *Psychology of Music*, 42(6), 869-887. <https://doi.org/10.1177/03057356145485>
- Folstein, M. F., Robins, L. N., & Helzer, J. E. (1983). The mini-mental state examination. *Archives of General Psychiatry*, 40(7), 812-812.
<https://doi.org/10.1001/archpsyc.1983.01790060110016>
- Hackert, M. Q., van Exel, J., & Brouwer, W. B. (2020). Well-being of Older People (WOOP): Quantitative validation of a new outcome measure for use in economic evaluations. *Social Science & Medicine*, 259, 113109.
<https://doi.org/10.1016/j.socscimed.2020.113109>
- Hsu, M. H., Flowerdew, R., Parker, M., Fachner, J., & Odell-Miller, H. (2015). Individual music therapy for managing neuropsychiatric symptoms for people with dementia and their carers: a cluster randomized controlled feasibility study. *BMC Geriatrics*, 15, 1-19. <https://doi.org/10.1186/s12877-015-0082-4>
- James, C. E., Müller, D. M., Müller, C. A. H., Van De Looij, Y., Altenmüller, E., Kliegel, M., Van De Ville, D., & Marie, D. (2024). Randomized controlled trials of non-pharmacological interventions for healthy seniors: Effects on cognitive decline, brain

- plasticity and activities of daily living-A 23-year scoping review. *Heliyon*, 10(9), e26674. <https://doi.org/10.1016/j.heliyon.2024.e26674>
- Jang, S., & Kunde, L. (2021). A systematic review of music therapy interventions used to address emotional needs of older adults. *The Arts in Psychotherapy*, 76, 1–11. <https://doi.org/10.1016/j.aip.2021.101842>
- Ito, E., Nouchi, R., Dinét, J., Cheng, C. H., & Husebø, B. S. (2022). The effect of music-based intervention on general cognitive and executive functions, and episodic memory in people with mild cognitive impairment and dementia: a systematic review and meta-analysis of recent randomized controlled trials. In *Healthcare* (Vol. 10, No. 8, p. 1462). MDPI. <https://doi.org/10.3390/healthcare10081462>
- Li, C., Kumar, A. P., Wankhar, D., Kuppusamy, M., & Govindasamy, K. (2025). Effect of music therapy on sleep quality in elderly: A systematic review and meta-analysis. *PLOS ONE*, 20(11), e0334356. <https://doi.org/10.1371/journal.pone.0334356>
- Linn, B. S., Linn, M. W., & Gurel, L. (1968). Cumulative illness rating scale. *Journal of the American Geriatrics Society*, 16(5), 622–626. <https://doi.org/10.1111/j.1532-5415.1968.tb02103.x>
- Ma, G., & Ma, X. (2023). Music intervention for older adults: Evidence map of systematic reviews. *Medicine*, 102(48), e36016. <https://doi.org/10.1097/MD.00000000000036016>
- Machado, T. H., Fichman, H. C., Santos, E. L., Carvalho, V. A., Fialho, P. P., Koenig, A. M., & Caramelli, P. (2009). Normative data for healthy elderly on the phonemic verbal fluency task-FAS. *Dementia & Neuropsychologia*, 3(1), 55-60. <https://doi.org/10.1590/S1980-57642009DN30100011>
- Malekmohammadi, A., & Cheng, G. (2025). Music familiarization elicits functional connectivity between right frontal/temporal and parietal areas in the theta and alpha bands. *Brain Topography*, 38(1), 2. <https://doi.org/10.1007/s10548-024-01081-z>

- Mangiacotti, A. M., A., Cipriani, G., Ward, E., Franco, F., & Biasutti, M. (2022). Development and validation of the Music Cognitive Test: a music-based cognitive screening test. *Psychology of Music*, 030573562211008. <https://doi.org/10.1177/03057356221100851>
- Mangiacotti, A. M., A., Hsu, M. H., Barone, C., Van Puyvelde, M., Zandonà, A., Gabai, G., Biasutti, M., & Franco, F. (2024a). Effects of one-to-one music therapy in older adults with cognitive impairment: A randomized controlled trial. *Psychology and Aging*, 39(8), 960. <https://doi.org/10.1037/pag0000861>
- Mangiacotti, A. M. A., Masotti, A., & Biasutti, M. (2024b). The impact of community music therapy intervention in older adults' care home communities: Assessing community music therapists' views. *International Journal of Community Music*, 17(1), 37-60. https://doi.org/10.1386/ijcm_00096_1
- Mondini, S., Mapelli, D., Vestri, A., & Bisiacchi, P. S. (2011). Esame neuropsicologico breve 2: Una batteria di test per lo screening neuropsicologico. [A battery of tests for neuropsychological screening]. Raffaello Cortina Editore. <https://hdl.handle.net/11577/2429698>
- Nasreddine, Z. S., Phillips, N. A., Bäckström, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- National Academy for Social Prescribing. (2024). *The impact of social prescribing on health service use and costs: Examples of local evaluations in practice* [Report]. https://socialprescribingacademy.org.uk/media/ibtdvgn0/nasp_sp_impactonservice_no_v24.pdf

- Nucci, M., Mapelli, D., & Mondini, S. (2012). Cognitive reserve index questionnaire (CRIq): a new instrument for measuring cognitive reserve. *Aging Clinical and Experimental Research*, 24(3), 218–226. <https://doi.org/10.1007/BF03654795>
- Oosterhuis, E. J., Slade, K., May, P. J. C., & Nuttall, H. E. (2023). Toward an understanding of healthy cognitive aging: The importance of lifestyle in cognitive reserve and the scaffolding theory of aging and cognition. *The Journals of Gerontology: Series B*, 78(5), 777–788. <https://doi.org/10.1093/geronb/gbac197>
- Pearce, E., Launay, J., Machin, A., & Dunbar, R. I. M. (2016). Is group singing special? Health, well-being and social bonds in community-based adult education classes. *Journal of Community & Applied Social Psychology*, 26(6), 518–533. <https://doi.org/10.1002/casp.2278>
- Panico, F., Sagliano, L., Magliacano, A., Santangelo, G., & Trojano, L. (2023). The relationship between cognitive reserve and cognition in healthy adults: a systematic review. *Current Psychology*, 42(28), 24751–24763. <https://doi.org/10.1007/s12144-022-03523-y>
- Raja, S., Barry, C., Upadhyay, R., Alash, R., O’Raghallaigh, M., Hayes, R., & Romero-Ortuno, R. (2025). Harmonious ageing: A narrative review of music therapy in the biomedical literature. *European Geriatric Medicine*, 16(2), 461–472. <https://doi.org/10.1007/s41999-024-01146-z>
- Robb, S. L., Springs, S., Edwards, E., Golden, T. L., Johnson, J. K., Burns, D. S., ... & Perkins, S. M. (2025). Reporting Guidelines for Music-based Intervention: an update and validation study. *Frontiers in psychology*, 16, 1551920. <https://doi.org/10.3389/fpsyg.2025.1551920>
- Sun, X., Liu, W., Gao, Y., Qin, L., Feng, H., Tan, H., Chen, Q., Peng, L., & Wu, I. X. Y. (2023). Comparative effectiveness of non-pharmacological interventions for frailty: a

- systematic review and network meta-analysis. *Age and Aging*, 52(2).
<https://doi.org/10.1093/aging/afad004>
- Tarrant, M., Lamont, R. A., Carter, M., Dean, S. G., Spicer, S., Sanders, A., & Calitri, R. (2021). Measurement of shared social identity in singing groups for people with aphasia. *Frontiers in Psychology*, 12, 669899.
<https://doi.org/10.3389/fpsyg.2021.669899>
- Ward, E. V., Isac, A., Donnelly, M., Van Puyvelde, M., & Franco, F. (2021). Memory improvement in aging as a function of exposure to mood-matching music. *Acta Psychologica*, 212, 103206. <https://doi.org/10.1016/j.actpsy.2020.103206>
- Wang, Y. Y., Wang, X. X., Chen, L., Liu, Y., & Li, Y. R. (2023a). A systematic review and network meta-analysis comparing various non-pharmacological treatments for older people with mild cognitive impairment. *Asian Journal of Psychiatry*, 103635.
<https://doi.org/10.1016/j.ajp.2023.103635>
- Wang, M., Wu, J., & Yan, H. (2023b). Effect of music therapy on older adults with depression: A systematic review and meta-analysis. *Complementary Therapies in Clinical Practice*, 53, 101809. <https://doi.org/10.1016/j.ctcp.2023.101809>
- WFMT - World Federation of Music Therapy. (2011). *What is music therapy?* World Federation of Music Therapy. <https://www.wfmt.info/about>
- World Health Organization. (2024). Ageing and health. <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health>
- Xu, B., Sui, Y., Zhu, C., Yang, X., Zhou, J., Li, L., Ren, L., & Wang, X. (2017). Music intervention on cognitive dysfunction in healthy older adults: A systematic review and meta-analysis. *Neurological Sciences*, 38(6), 983–992.
<https://doi.org/10.1007/s10072-017-2878-9>

Zhang, Y., Li, R., Zhu, H., & Chen, J. (2024). Music therapy for adults aged sixty and older: An umbrella review of systematic reviews. *Archives of Gerontology and Geriatrics*, 118, 105180. <https://doi.org/10.1016/j.archger.2024.105180>

Zhou, M., Wei, C., Xie, X., Wang, Z., Chen, L., & Zhang, X. (2025). Effect of music therapy on cognitive function among older adults: A systematic review and meta-analysis of randomized controlled trials. *The Gerontologist*, 65(6), gnaf106. <https://doi.org/10.1093/geront/gnaf106>

Footnotes

¹ Large-group MT sessions comprised at least 15 individuals. Twenty-one participants across three large MT groups took part in the study, but of these, 12 completed the study in full and were entered into the analysis.

Table 1. *Characteristics of the final sample at baseline*

Variable	One-to-one (<i>n</i> =17)	Small group (<i>n</i> =17)	Large group (<i>n</i> =12)	Control (<i>n</i> =16)	<i>p</i>	ηp^2
Age (years)	79.71 (6.08) [70-92]	82.00 (7.17) [66-93]	79.83 (11.58) [53-101]	74.56 (7.47) [61-88]	.067	.115
Sex					.643	.164*
Female	12 (70.6%)	14 (82.4%)	10 (83.3%)	14 (87.5%)		
Male	5 (29.4%)	3 (17.6%)	2 (16.7%)	2 (12.5%)		
MMSE	28.82 (1.07) [27-30]	27.76 (1.39) [26-30]	27.92 (1.62) [25-30]	28.50 (1.27) [26-30]	.093	.104
CRIq	140.88 (22.81) [85-175]	135.41 (16.67) [96-160]	119.92 (20.23) [90-154]	129.31 (23.22) [95-166]	.061	.118
CIRS	7.59 (4.64) [0-18]	9.06 (5.33) [0-15]	6.00 (3.02) [0-11]	6.13 (4.50) [0-18]	.212	.074
CIRS_S	0.54 (0.33) [0-1.29]	0.65 (0.38) [0-1.07]	0.43 (0.22) [0-0.79]	0.44 (0.32) [0-1.29]	.212	.074
CIRS_C	0.41 (0.62) [0-20]	0.59 (0.80) [0-2]	0.17 (0.39) [0-1]	0.50 (0.73) [0-2]	.399	.049
MCT	48.82 (2.43) [43-52]	46.76 (2.82) [43-52]	47.08 (3.50) [38-51]	48.69 (2.92) [41-52]	.101	.101
TMT-A	48.24 (20.24) [22.00-112.51]	58.04 (22.72) [32.21-99.50]	49.58 (14.01) [30.00-72.16]	48.28 (16.78) [23.87-77.52]	.394	.050
TFT	27.18 (6.50) [15-38]	23.06 (10.01) [7-37]	22.50 (6.79) [15-37]	24.88 (5.98) [15-36]	.316	.059
TFT_V	6.53 (4.54) [0-16]	5.06 (3.32) [0-12]	5.75 (1.60) [3-8]	1.81 (1.72) [0-6]	.001	.262
TFT_R	1.29 (1.53) [0-5]	2.12 (1.99) [0-7]	1.50 (1.51) [0-4]	1.44 (1.96) [0-6]	.555	.035
TFT_S	0.59 (0.795) [0-2]	0.35 (0.71) [0-2]	.33 (0.77) [0-2]	0.50 (0.89) [0-3]	.780	.018
TFT_A	0.18 (0.39) [0-1]	0.18 (0.39) [0-1]	0.42 (0.51) [0-1]	0.69 (1.13) [0-3]	.116	.096
VF	19.88 (6.47) [1-31]	18.00 (6.18) [9-33]	14.92 (3.77) [10-21]	15.38 (3.78) [9-21]	.041	.131
GDS	2.24 (1.89) [0-7]	2.88 (2.52) [0-8]	3.58 (3.87) [0-12]	2.50 (2.22) [0-7]	.564	.034
WOOP	38.59 (4.00) [32-44]	36.00 (6.02) [25-44]	36.17 (6.03) [26-44]	37.50 (5.22) [28-45]	.479	.042
PSQI	9.94 (4.86) [3-21]	15.06 (5.65) [7-32]	15.67 (7.02) [2-33]	12.25 (6.01) [2-31]	.030	.141

Note. MMSE = Mini Mental State Exam; CRIq = Cognitive Reserve Index questionnaire; CIRS = Cumulative Index score for physical health; CIRS_S = CIRS Severity Index; CIRS_C = CIRS Comorbidity Index; MCT = Music Cognitive Test; TMT-A = Trail Making Test parts A; TFT = Tangle Figure Test; VF = Verbal Fluency; GDS = Geriatric Depression Scale (short version); WOOP = Well-being of Older People Scale; PSQI = Pittsburgh Sleep Quality Index. For sex, sample sizes and percentages are reported. For all other variables, mean, standard deviation (in parentheses), and range (in brackets) are reported.

*Cramér's V

Table 2. ANOVA outcomes for each measure as a function of Time (pre-post intervention) and Group (one-to-one, small-group, large-group, control)

Measure	Source of Variance	SS	df	F	p	η_p^2
MCT	Time	9.807	1	3.847	.055	.062
	Group	34.544	3	0.759	.522	.038
	Time $\times$ Group	52.879	3	6.915	<.001*	.263
	Error	147.839	58	—	—	—
TMT-A	Time	439.394	1	4.265	.043*	.068
	Group	1,882.709	3	0.897	.448	.044
	Time $\times$ Group	1,041.031	3	3.368	.024*	.148
	Error	5,976.021	58	—	—	—
TFT	Time	457.648	1	33.996	<.001*	.37
	Group	595.956	3	1.596	.201	.076
	Time $\times$ Group	125.593	3	3.11	.033*	.139
	Error	780.794	58	—	—	—
VF	Time	57.226	1	4.307	.042*	.069
	Group	701.743	3	4.718	.005*	.196
	Time $\times$ Group	24.077	3	0.604	.615	.030
	Error	770.640	58	—	—	—
TFT_V	Time	1.071	1	0.158	.692	.003
	Group	145.757	3	2.688	.055	.122
	Time $\times$ Group	68.725	3	3.386	.024*	.149
	Error	392.412	58	—	—	—
TFT_R	Time	0.012	1	0.006	.938	.000
	Group	18.648	3	1.518	.219	.073
	Time $\times$ Group	15.195	3	2.597	.061	.118
	Error	113.104	58	—	—	—
TFT_S	Time	2.147	1	6.125	.016*	.096
	Group	0.234	3	0.167	.918	.009
	Time $\times$ Group	0.836	3	0.795	.502	.039
	Error	20.333	58	—	—	—
TFT_A	Time	0.283	1	0.782	.380	.013
	Group	1.479	3	0.870	.462	.043
	Time $\times$ Group	2.741	3	2.527	.066	.116
	Error	20.969	58	—	—	—
GDS	Time	35.695	1	19.121	<.001*	.248
	Group	12.728	3	0.471	.703	.024
	Time $\times$ Group	5.695	3	1.017	.392	.05
	Error	108.273	58	—	—	—
WOOP	Time	2.637	1	0.528	.478	.009
	Group	194.528	3	1.387	.256	.067
	Time $\times$ Group	12.727	3	0.85	.472	.042
	Error	289.627	58	—	—	—
PSQI	Time	9.050	1	0.795	.376	.014
	Group	449.301	3	2.152	.103	.1
	Time $\times$ Group	25.892	3	0.758	.522	.038
	Error	660.051	58	—	—	—

Note. MD = Mean difference; MCT = Music Cognitive Test; TMT-A = Trial Making Test-A; TFT = Tangle Figure Test; TFT_V = Tangle Figure Test Visual Errors = TFT_R = Tangle Figure Test Repetition Errors; TFT_S = Tangle Figure Test Semantic Errors; TFT_A = Tangle Figure Test Anomia; VF = Verbal Fluency; GDS = Geriatric Depression Scale (short version). WOOP = Well-being of Older People Scale. PSQI = Pittsburgh Sleep Quality Index. * = $p < .05$

Table 3. Bonferroni-corrected pairwise comparisons of outcome measures pre-post intervention

Intervention	Measure	Baseline mean (SD)	Post mean (SD)	<i>p</i>	η^2	MD [95% CI]
One-to-one MT (<i>n</i>=17)	MCT	48.82 (2.43)	49.18 (1.43)	.522	.007	0.35 [-1.45, 0.74]
	TMT-A	48.24 (20.24)	39.71 (13.56)	.017*	.094	8.53 [1.57, 15.51]
	TFT	27.18 (6.50)	32.47 (8.19)	<.001*	.234	5.29 [2.78, 7.81]
	TFT_V	6.53 (4.54)	5.47 (3.41)	.240	.024	1.06 [-0.73, 2.85]
	TFT_R	1.29 (1.53)	1.71 (1.53)	.394	.013	-0.41 [-1.37, 0.55]
	TFT_S	0.59 (0.80)	0.06 (0.24)	.012*	.105	0.53 [0.12, 0.94]
	TFT_A	0.18 (0.39)	0.41 (0.71)	.259	.022	-0.24 [-0.65, 0.18]
	VF	19.88 (6.47)	21.82 (7.64)	.126	.040	-1.94 [-4.44, 0.56]
	GDS	2.24 (1.89)	1.29 (1.36)	.049*	.065	0.94 [0.00, 1.88]
	WOOP	38.59 (4.00)	39.35 (4.36)	.323	.017	-0.76 [-2.30, 0.77]
	PSQI	9.94 (4.87)	11.53 (5.20)	.175	.031	-1.59 [-3.91, 0.73]
Small-group MT (<i>n</i>=17)	MCT	46.76 (2.82)	48.71 (2.69)	<.001*	.178	1.94 [0.85, 3.04]
	TMT-A	58.04 (22.72)	49.95 (21.55)	.024*	.085	8.09 [1.12, 15.06]
	TFT	23.06 (10.01)	25.29 (10.18)	.081	.052	2.24 [-4.75, 0.28]
	TFT_V	5.06 (3.33)	4.76 (3.77)	.743	.002	0.29 [-1.49, 2.08]
	TFT_R	2.12 (2.00)	1.65 (1.90)	.330	.016	0.47 [-0.49, 1.43]
	TFT_S	0.35 (0.70)	0.24 (0.56)	.565	.006	0.12 [-0.29, 0.52]
	TFT_A	0.18 (0.39)	0.12 (0.33)	.776	.001	0.06 [-0.35, 0.47]
	VF	18.00 (6.17)	20.53 (6.28)	.048*	.066	-2.52 [-5.03, -0.03]
	GDS	2.88 (2.52)	1.88 (2.18)	.037*	.073	1.00 [0.062, 1.94]
	WOOP	36.00 (6.02)	36.71 (4.97)	.361	.014	-0.71 [-2.24, 0.83]
	PSQI	15.06 (5.65)	15.53 (6.59)	.686	.003	-0.47 [-2.78, 1.85]
Large-group MT (<i>n</i>=12)	MCT	47.08 (3.50)	48.50 (2.15)	.034*	.075	1.42 [0.11, 2.72]
	TMT-A	49.58 (14.01)	45.48 (20.78)	.326	.017	4.11 [-4.19, 12.40]
	TFT	22.50 (6.79)	29.00 (9.30)	<.001*	.245	6.51 [3.50, 9.50]
	TFT_V	5.75 (1.60)	5.17 (4.04)	.585	.005	0.58 [-1.54, 2.71]
	TFT_R	1.50 (1.51)	2.42 (2.39)	.113	.043	-0.92 [-2.06, 0.23]
	TFT_S	0.33 (0.78)	0.17 (0.39)	.493	.008	0.17 [-0.32, 0.65]
	TFT_A	0.42 (0.52)	0.42 (1.16)	1.000	.000	0.00 [-0.49, 0.49]
	VF	14.92 (3.72)	15.25 (3.86)	.824	.001	-.333 [-3.31, 2.64]
	GDS	3.58 (3.87)	1.75 (2.49)	.002*	.157	1.83 [0.72, 2.95]
	WOOP	36.17 (6.03)	35.25 (5.72)	.319	.017	0.92 [-0.91, 2.74]
	PSQI	15.67 (7.02)	14.67 (8.27)	.471	.009	1.00 [-1.76, 3.76]
Control (<i>n</i>=16)	MCT	48.69 (2.92)	47.25 (4.80)	.014*	.101	-1.44 [-2.57, -0.31]
	TMT-A	48.28 (16.78)	53.79 (26.12)	.130	.039	-5.51 [-12.70, 1.67]
	TFT	24.88 (5.98)	26.38 (8.14)	.252	.023	1.51 [-4.10, 1.10]
	TFT_V	1.81 (1.72)	4.50 (4.23)	.005*	.128	-2.69 [-4.53, -0.85]
	TFT_R	1.44 (1.97)	0.50 (0.73)	.063	.059	0.94 [-0.05, 1.93]
	TFT_S	0.50 (0.89)	0.25 (0.45)	.237	.024	0.25 [-0.17, 0.67]
	TFT_A	0.69 (1.14)	0.13 (0.34)	.010*	.108	0.56 [0.14, 0.99]
	VF	15.38 (3.77)	16.06 (4.23)	.596	.005	-0.68 [-3.26, 1.89]
	GDS	2.50 (2.22)	1.94 (1.98)	.249	.023	0.56 [-0.40, 1.53]
	WOOP	37.50 (5.22)	38.13 (4.46)	.432	.011	-0.63 [-2.21, 0.96]
	PSQI	12.25 (6.01)	13.37 (7.49)	.349	.015	-1.13 [-3.51, 1.26]

Note. MT = music therapy; MD = Mean difference; MCT = Music Cognitive Test; TMT-A = Trial Making Test-A; TFT = Tangle Figure Test; TFT_V = Tangle Figure Test Visual Errors; TFT_R = Tangle Figure Test Repetition Errors; TFT_S = Tangle Figure Test Semantic Errors; TFT_A = Tangle Figure Test Anomia; VF = Verbal Fluency; GDS = Geriatric Depression Scale (short version). WOOP = Well-being of Older People Scale. PSQI = Pittsburgh Sleep Quality Index. Recommendations from Tian et al. (2018) were considered for post-hoc comparison when omnibus tests were not significant.

* = $p < .05$

Figure 1. *Description of the Randomized Controlled Trial participant recruitment process*

Note. MT = music therapy

Alt Text: Flowchart showing participant recruitment, enrolment, randomization, and analysis:

413 approached, 148 interested, 86 included, then randomized into four groups, with 62 participants ultimately analysed after withdrawals and exclusions.

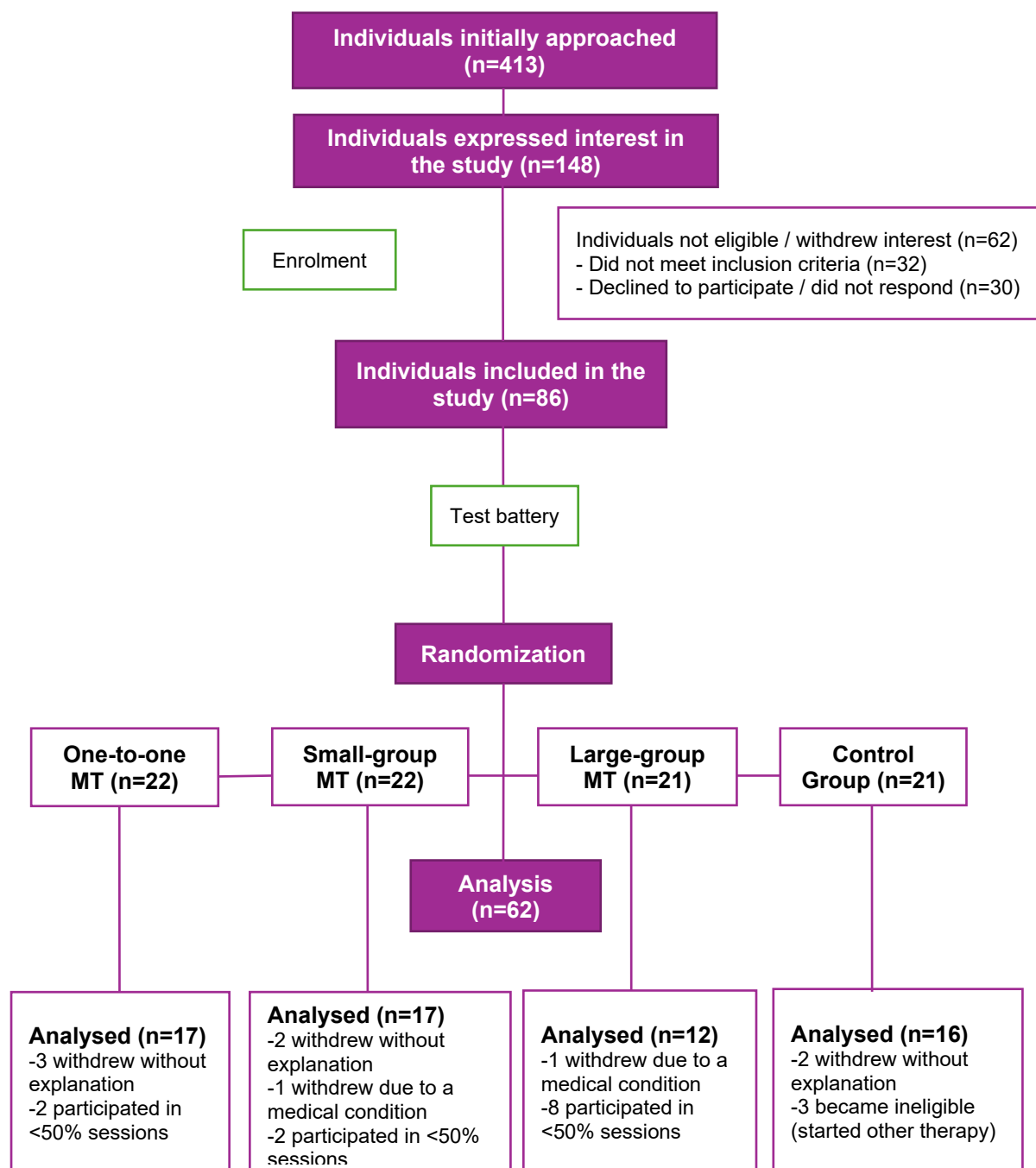


Figure 2. *Performance pre-post intervention: (A) Effects on cognitive tests; (B) Effects on well-being tests.*

Note. * = $p < 0.05$

Alt Text: Multi-panel line graphs showing pre–post changes across the four groups on the cognitive and well-being tests, showing overall improvements in intervention groups.

