Effects of One-to-one Music Therapy In Older Adults With Cognitive Impairment – An RCT

Abstract [194/250w]

This study investigated the effects of music therapy (MT), a non-pharmacological therapy, on cognitive, behavioural, and physiological outcomes in older adults with mild to moderate cognitive decline residing in care home settings. A Randomized Controlled Trial (RCT) design was employed, with 42 care home residents (mean age 86.25 years) randomly assigned to either a one-to-one 16-week MT intervention or an active control group receiving story telling (ST). Experimental and control activities were matched on key aspects, and groups were equivalent at baseline concerning demographic factors, general health, cognitive-behavioural characteristics, and cognitive reserve levels. Pre/post-intervention neuropsychological and behavioural measures were collected, alongside saliva samples for cortisol/dehydroepiandrosterone ratio and respiratory sinus arrhythmia analysis as indicators of overall stress and autonomic regulation. The MT group exhibited benefits in cognitive, behavioural, and physiological domains, suggesting potential advantages in maintaining cognitive functioning and reducing neuropsychiatric symptoms. Biomarkers indicated possible mechanisms underlying the effectiveness of MT. The sixteen sessions of one-to-one interactive MT - delivered over five months - had a positive impact on older adults with mild to moderate cognitive decline living in care home settings. The implications of these findings for healthy aging and suggestions for future research are discussed.

Public Significance Statement

1. Music therapy demonstrated positive effects on cognitive, behavioural, and mood symptoms in older adults with cognitive decline.
2. Music therapy showed potential in reducing stress-related biomarkers, such as salivary cortisol/DHEA, and improving cardiorespiratory frequency.
3. While the findings suggest promising outcomes, further research is needed to explore the role of music therapy in the broader context of cognitive health.

Keywords: Music therapy, older adults, cognitive impairment, neuropsychiatric symptoms, hormones

Introduction

The number of older adults living with neurocognitive disorders is increasing and is estimated to have further risen by 2050 (Gauthier et al., 2021; WHO, 2017). Neurocognitive disorders are defined as conditions affecting cognitive functions (e.g., memory, language) that are associated with an increased risk of behavioural changes (e.g., apathy, agitation, depression) (Dillon et al., 2013). Cognitive and behavioural changes can arise from diverse factors (Savva et al., 2009). Recognizing the complexity of their origins highlights the potential for incorporating behavioural interventions alongside medical considerations. Different pharmacological interventions are used to improve cognitive symptoms, but with limited effects on mitigating the related behavioural ones (Dyer et al., 2018; Huang et al., 2023). In addition, psychotropic drugs are used to obtain more stable effects, resulting in increased costs of pharmacological therapy (Skoldunger et al., 2013). With ageing, there is a higher sensitivity to drugs and susceptibility to adverse drug reactions (Shi et al., 2008), hence the need for non-pharmacological and less expensive interventions to improve the quality of life in this population. Policymakers are sensitive to issues associated with mental health challenges in ageing and are starting to embrace social prescribing, an approach that connects people to non-medical resources and activities in their community to address their social, emotional, and practical needs that can significantly impact their well-being and overall health such as, arts and crafts, music, physical exercises etc (Costa, et al., 2021; Drinkwater et al., 2019). These activities are person-centred focusing on a holistic view of the person (Grover et al., 2023). A wealth of research has flourished to study non-pharmacological intervention to minimize behavioural disturbances and improve cognitive functioning and quality of life in people with neurocognitive disorders (e.g., Batubara et al., 2023; Hudak et al, 2019; Laver et al., 2016; Wang et al., 2023b;)

Music therapy (MT) is a non-pharmacological, person-centred, psychological intervention that relies on the systematic and evidence-based use of musical interventions by credentialed professionals within a therapeutic relationship. MT is employed to address the physical, emotional, cognitive, and social needs of individuals, including those with cognitive decline (Lin et al., 2023; Moreno-Morales et al., 2020; Wang et al., 2023b), offering promising benefits. These benefits include enhancement in social functioning (e.g., communication, cooperation, empathy), cognitive (e.g., attention, working memory), mood (lowering the risk of depression), cerebral plasticity, and reducing behavioural symptoms (Altenmüller & Schlaug, 2015; Hsu et al., 2015; Suzuki et al., 2004; Wang et al., 2023b; Yao et al., 2023; Zhang et al., 2017). However, due to the nature of MT, which addresses clinical and psychological aspects,

there has been a shift towards a quantitative approach rather than a qualitative one to evaluate the potential benefits of the therapy. More RCTs with high methodological reporting quality are being developed (de Witte et al., 2022; Fusar-Poli et al., 2017; Leggieri et al., 2019), implementing structured theoretical models, and incorporating appropriate outcome measures (Moreno-Morales et al., 2020; Wang et al., 2023b). These efforts are crucial for enhancing the scientific evidence supporting the effectiveness of MT. It was suggested that studies should involve people with neurocognitive disorders who are older than 80 using both psychosocial and biochemical measures (Chu et al., 2013). Moreover, literature reviews (Li et al., 2019; Moreno-Morale, et al., 2020; Wang et al., 2023a; Xu et al.,2017) argued the need for enhanced generalizability of certain findings to make clear whether MT is an effective treatment to improve cognitive abilities and behavioural symptoms, especially in those suffering from mild-moderate cognitive decline. Besides, it was suggested that future studies should investigate the combination of endocrinological, behavioural and cognitive measures to tailor MT effects in people with cognitive decline (Suzuki et al., 2004).

Based on these premises, we investigated effects of a 4-month, one-to-one MT intervention compared to a matched Story Telling intervention (ST henceforth) in older adults with mild to moderate cognitive decline living in care homes. We applied an in-depth investigation, using cognitive, behavioural, endocrinological, and physiological measures to increase our insight into potential physiological and neuropsychological effects of MT, as recommended in the literature (Chu et al., 2013; Li et al, 2019; Suzuki et al., 2004; Xu et al., 2017).

Besides standard neuropsychological tests, our primary outcome measures for the cognitive and behavioural domain were respectively: [i] the Music Cognition Test (MCT, Mangiacotti et al., 2022), and the [ii] Cornell Scale for Depression in Dementia (CSDD, Alexopoulos et al., 1988). Two secondary outcomes were adopted [i] hormonal cortisol/dehydroepiandrosterone (DHEA) ratio, and [ii] Respiratory Sinus Arrhythmia (RSA). The MCT was chosen as an outcome measure because it is sensitive to the specific skills targeted by music training. It covers the same main cognitive components as the MOCA (Nasreddine et al., 2005) and MMSE (Folstein et al., 1975), using musical task items with which it is strongly correlated. This sensitivity allows for a more accurate assessment of the cognitive functions stimulated by near transfer effects of music training, where improvements are observed in tasks closely related to the training activities. In contrast, far transfer effects would generalize to more distantly related cognitive functions. The CSDD serves as the primary behavioral measure, evaluating signs of major depression in individuals

with cognitive decline from both participant and caregiver perspectives. Salivary cortisol and DHEA were analysed as endocrinological measures to assess the intervention's long-term effects. As endocrinological measures, salivary cortisol and DHEA were analysed to measure long-term effects of the intervention. Cortisol is an adrenocortical hormone involved in the stress response and is widely used as a stress index (Fancourt et al., 2016; Almeida et al., 2020). Sustained stress is related to high levels of circulating cortisol, which can lead to behavioural and psychological disorders (e.g., mood disorder or anxiety syndrome) and can be associated with a physiological stressor (Bassett et al., 1987; Stahl & Dörner, 1982). Stress appears to be involved in the progression of dementia disease (Gilsanz et al., 2019), with high levels of cortisol causing damage to neurons, particularly in the hippocampus (Hill & Spencer-Segal, 2021; Kimonides et al., 1999; Lupien et al., 1998), a component of the brain that plays an important role in the consolidation of memories (Chu et al., 2014; de La Rubia Ortí et al., 2018; Takahashi & Matsushita, 2006). DHEA and its sulphated (DHEA-s) are one of the most abundant naturally occurring adrenal steroids that show a marked age-related decline in humans (van Niekerk et al., 2001; Tannenbaum et al., 2003). The major biological action of DHEA in the brain involves neuroprotection (Kurata et al., 2004; Li et al., 2001), neurogenesis and neuronal preservation (Karishma & Herbert, 2002; Sakr et al., 2014), and antioxidative and anti-inflammatory responses (Chen & Parker, 2004; Tamagno et al., 2003). DHEA-s could play a protective role during stress as an antagonist to the neuro-toxic consequences of elevated cortisol (Morgan et al., 2004). However, with chronic stress, the capacity to produce DHEA-s is reduced (Lennartsson, Theorell, Kushnir, et al., 2013), and DHEA-s levels are lower (Lennartsson et al., 2013). Notably, in adults, positive social interactions may contribute to a more favourable physiological response through increased DHEA-S levels (Polenick et al., 2021). All these effects explain the associations of DHEA with well-being, mood and cognitive functions, e.g., enhancing memory and executive functions (de Menezes et al., 2016; Hildreth et al., 2013; Zheng, 2009). However, the main endocrinological measure selected for the present study is the cortisol/DHEA ratio, which denotes the balance between catabolic and anabolic activity and can be used as an indicator of adverse chronic stress (Theorell et al., 2021). A relationship between these two biomarkers is highlighted (Frye & Lacey 1999): in high-stress situations, as cortisol increases, DHEA decreases, leading to an increased cortisol/DHEA ratio. In relaxed conditions, DHEA increases while cortisol levels reduce, decreasing the cortisol/DHEA ratio. In older adults, this ratio appears related to an individual's cognitive level, with higher ratios associated with the risk of cognitive impairment (Kalmijn et al., 1998; Yiallouris et al., 2019) and less brain plasticity (Aoki et al., 2020;

Ji et al., 2021; Jin et al., 2016). Few studies have investigated the effects of music-based activities on the cortisol/DHEA ratio in an ageing adult population. However, some encouraging findings suggested a positive effect of musical activities with decreased ratios in both an ageing (Theorell et al., 2021) and non-ageing population (Bittman et al., 2001; Fancourt & Williamon, 2016). Therefore, MT studies that analyse Cortisol/DHEA in an ageing population with cognitive impairment are needed to consolidate these initial suggestions.

To enhance the precision and depth of our study's investigation into the physiological therapeutic mechanisms among interventions, autonomic regulation, emotional states, and cognitive functions, we chose Respiratory Sinus Arrhythmia (RSA) as a key physiological metric. RSA is a non-invasively measured output of the interaction between the respiration and the heart rate, thought to reflect the parasympathetic involvement of the autonomic nervous system. RSA illustrates how heart rate (HR) changes during the respiration cycle, i.e., HR accelerates during inspiration due to sympathetic activity and decelerates during expiration due to parasympathetic activity (Berston et al., 1997; Grossman et al., 1990). The variability between heartbeats during inspiration and deceleration makes it possible to obtain heart rate variability (HRV), an index of autonomic activity (Berntson et al., 1997). Within HRV, RSA reflects the parasympathetic-inhibitory impact on the heart, mediated by the vagus nerve (Berston et al., 1997; Thayer & Lane, 2007). Research suggested that RSA is associated with socio-emotional responding (Butler et al., 2006; Frazier et al., 2004). Higher RSA reflects greater vagal control of the heart, associated with increased parasympathetic activity, that is thought to promote social communication (Patriquin et al., 2013). Increased parasympathetic activity of the heart is manifested in safe and relaxing situations. Meanwhile, increased sympathetic division of the nervous system is activated in response to stress situations and the flexible interaction between both branches allows the body to rapidly respond to these situations (Thayer & Lane, 2000, 2009; Van Puyvelde et al., 2014). Interestingly, theoretical models discussed how higher resting RSA is linked with enhanced cognitive performance (Thayer & Lane, 2000, 2009). Several studies demonstrated that increased resting parasympathetic regulation may be associated with better cognitive ability (Lü and Wang, 2018; Morgan et al., 2007; Staton et al., 2009), more adaptive cognitive functioning (Beauchaine, 2001), and increased performance in attention and working memory tasks (Hansen et al., 2003), as well as in executive control tasks (Mathewson et al. 2011). Based on these findings, we deemed RSA a good candidate for measuring short-term intervention effects in an older adult population with cognitive decline.

In summary, the research questions addressed by this study concerned whether there was a difference in the impact of a MT versus ST intervention on (i) the cognitive-neuropsychological, (ii) behavioural-wellbeing and (iii) psychophysiological domains, (i.e., cortisol/DHEA ratio, RSA) in older individuals with mild-moderate cognitive impairment.

Method

Transparency and Openness

We report how we determined our sample size and describe all manipulations, measures that were collected, RCT design, inclusion/exclusion criteria and statistical analyses as described in our pre-registration at ClinicalTrial.gov (n. <Anonymised>). The anonymised dataset on which the study results and conclusions are based are accessible at the <university> Figshare page (https://doi.org/10.22023/ANONYMISED). Data were analysed using SPSS, Version 25 and R. Due to the complexity of the training and copyright restrictions on some of the cognitive tests used, we were unable to post the full materials at the public link. However, information and materials related to the MT and ST activities and protocols are detailed in the supplementary material.

Study Design

An RCT design with a quantitative approach was used. The study was developed through previous collaborations between the study's researchers and <XXX>, a national charity providing care to older adults in the UK. The charity has a well-established relationship with several universities, as it supports research in the MT field. A total of five <XXX> care homes were selected within and outside Greater London to recruit a broadly representative sample, i.e., based not only on a metropolitan population (with multiple cultures/languages etc., particularly pronounced in London). Before starting the projects, ethical approval was granted by <XXX> University Psychology Research Ethics Committee, along with the Declaration of Helsinki principles. Before beginning the recruitment process, information about the study was provided in a meeting including care home staff and participants' family members. All participants gave written informed consent.

Participants

A priori power analysis was conduct using G*Power software version 3.1.9.7 (Erdfelder et al., 1996) with an effect size (f) = 0.25, $\alpha = 0.05$, Power (1-β) = 0.80 (any level over 0.80 is consider satisfactory; Aktas & Keskin, 2013), number of groups = 2 (MT & ST) and number of measures = 2 (pre-post design) yielding an overall N = 34 for

repeated measure ANOVA (within-between interaction). Thus, the sample size of N = 34 is adequate to test the study hypothesis. To account for possible drop out, an overrecruiting of 30% was considered. Timeframe of pre-post data collection and intervention was between Feb-Aug 2018.

Inclusion criteria were: i) aged 65 or older; ii) scored between 18-23 on the Mini-Mental State Examination (Folstein et al., 1975), considered an index of mild to moderate cognitive impairment (An & Liu, 2016; Tombaugh & McIntyre, 1992); iii) no significant hearing impairment that would negatively interfere with participation in the proposed activities. Exclusion criteria were: i) presence of severe motor deficits that would not allow participants to take part in the activities; ii) previous attendance in a cognitive training programme or MT programme within the last six months; iii) previous participation in a cognitive assessment within the last three months before the start of the pre-test phase.

Participants were randomly allocated by a researcher who was masked to treatment conditions to a MT (experimental group) or a ST intervention (active control group) using a computerised randomisation method. To minimise a possible drop-out rate with a consequent loss of power, a total of 50 participants were recruited (48% over the minimum requested sample size), of whom 42 completed the study (25% over the minimum requested sample size), 23 in the experimental group and 19 in the control group. The sample size of 42 participants included 95.24% individuals who identified as White and 4.76% who identified as Black-Caribbean in terms of ethnicity. All participants underwent a neuropsychological examination at baseline (within 2weeks prior starting the activities). No significant baseline differences were found between the experimental and control group as to screening demographic variables, MMSE, Cognitive Reserve (Nucci et al., 2012) and a battery of cognitive and behavioural tests (Table 1).

Table 1

Groups demographic data (n=42) of the participants who completed the study.

	Music Therapy group (*n* = 23)	**Story Telling group (*n* = 19)**	***p***	**effect size**
Age [range/mean(SD)]	62-95/84.86(7.37)	77-99/87.63(7.58)	.240	0.371[1]
Sex [female *n*(%)/male *n*(%)]	20 (87.00)/3(13.00)	15 (78.9)/4(21.1)	.481	0.107[2]
YoE [range/mean(SD)]	7-18/10.00(2.11)	7-19/10.26(2.88)	.743	0.035[1]
CIRS [range/mean(SD)]	1-19/7.43(5.11)	0-24/9.42(5.95)	.259	0.358[1]
CRIq [range/mean(SD)]	80-191/113.65(14.91)	82-189/115.21(16.59)	.753	0.101[1]
MMSE [range/mean(SD)]	18-24/19.57(1.75)	18-23/20.58(1.98)	.091	0.487[1]
MCT [range/mean(SD)]	21-46/32.52(6.90)	21-43/34.53(4.59)	.268	0.343[1]

MOCA [range/mean(SD)]	4-22/10.39(3.77)	4-24/12.89(5.09)	.075	0.559[1]
VFP [range/mean(SD)]	1-14.5/7.23(3.67)	0-18/7.89(4.81)	.629	0.154[1]
VFS [range/mean(SD)]	3-18/7.65(3.52)	1-19/7.68(4.55)	.980	0.007[1]
CDT [range/mean(SD)]	0-10/4.82(3.22)	2-10/5.60(3.46)	.462	0.233[1]
CSDD [range/mean(SD)]	1-20/7.17(4.30)	2-19/9.47(5.26)	.135	0.478[1]
BADLS [range/mean(SD)]	8-25/14.00(5.34)	2-30/15.21(8.46)	.576	0.171[1]
QoL [range/mean(SD)]	23-45.33/34.46(5.09)	24.33-42.33/32.92(5.18)	.342	0.297[1]
SWSL [range/mean(SD)]	12-35/24.17(6.11)	13-33/21.26(6.25)	.137	0.471[1]
NPI [range/mean(SD)]	0-84/ 12.65(16.83)	0-31/11.58(9.38)	.796	0.078[1]
NPId [range/mean(SD)]	0-29/4.04(5.41)	0-18/4.68(4.83)	.687	0.124[1]
MOSES-IV [range/mean(SD)]	8-21/11.52(3.75)	7-23/11.95(4.68)	.751	0.101[1]
MOSES-V [range/mean(SD)]	8-23/15.57(4.63)	11-24/17.05(4.10)	.277	0.338[1]

Note. t-test: Age; YoE, Years of Education; CIRS, Cumulative Illness Rating Scale; CRIq, Cognitive Reserve Index questionnaire; MMSE, MiniMental State Examination; MOCA, Montreal Cognitive Assessment; MCT, Music cognitive Test; VFP, verbal fluency phonetic; VFS: Verbal fluency Semantic; CDT, Clock Drawing Test; BADLS, Bristol Activities of Daily Living Scale; CSDD, Cornel Scale for Depression in Dementia; QoL, Quality of Life; SWSL, Satisfaction With Life Scale; NPI, Neuropsychiatric inventory; NPId, Neuropsychiatric inventory disruptiveness; MOSES, multidimensional Observation Scale for Elderly Subjects.

Chi-Square Test of independence: Sex

Effect Size: [1] Cohens d; [2] Cramer's V

Some participants withdrew from undertaking the Tangled Figure Test (TFT, Mondini et al., 2011) and Trail Making Test-A (TMT-A, Reitan & Wolfson, 1995). Withdrawal was due to non-cooperation by the participants in completing the tests. The total sample size of participants who completed the two tests is N = 31, which falls slightly below the minimum of 34 as indicated in the power analysis. Despite this minor loss of power, no significant difference was found at baseline within these two tests between the experimental and control groups (Table 2). The complete design is summarized in Figure 1.

Table 2

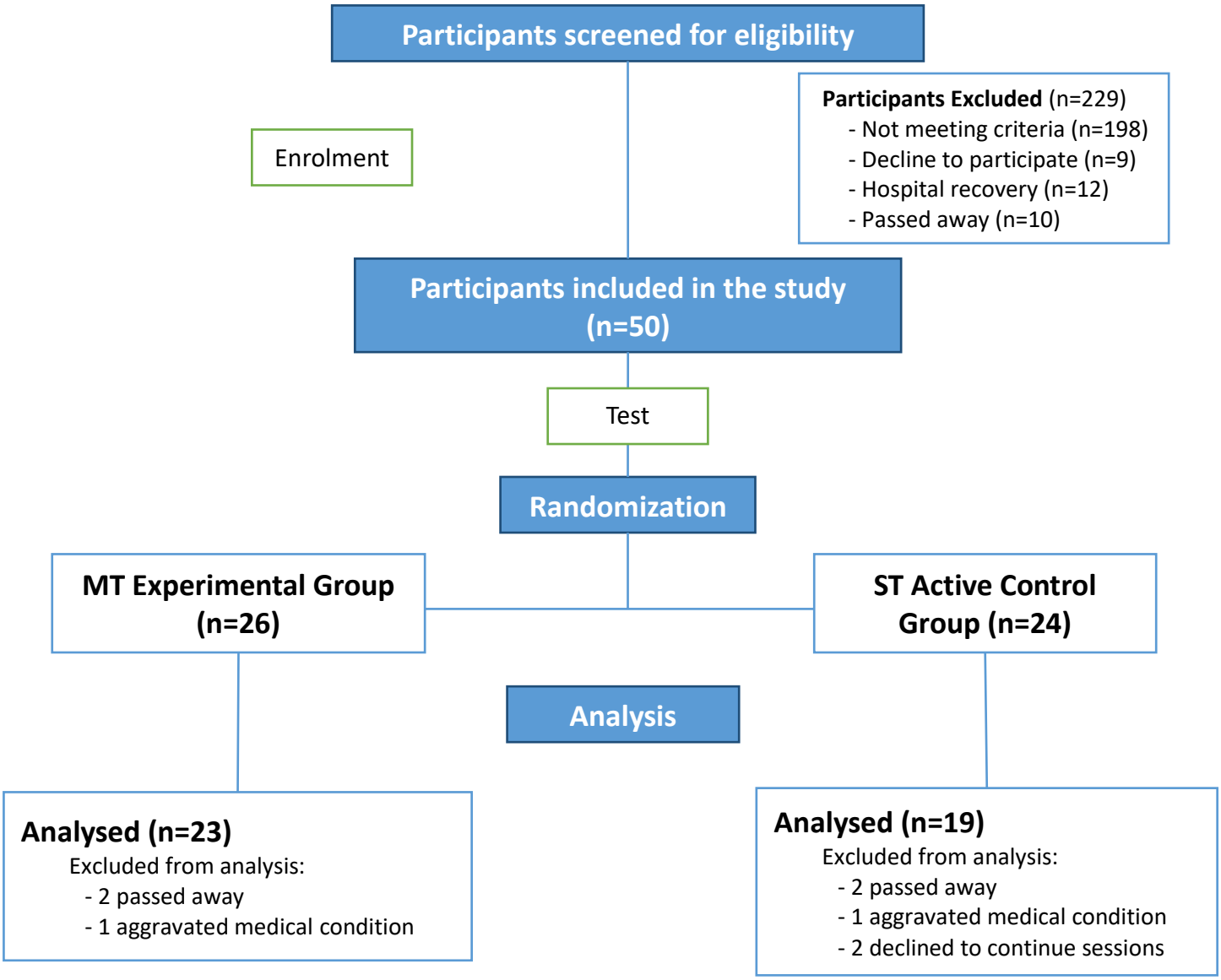
Groups demographic data of the participants who completed part of the cognitive test.

	n	**Music Therapy group**	**n**	**Story Telling group**	***p***	**Cohens' d**
TMT-A [range sec./mean(SD)]	18	50-720/187.61(149.21)	13	36-399.6/151.05(107.82)	.443	0.281
TFT [range/mean(SD)]	19	0-18/8.05(5.15)	13	0-26/12.46(7.96)	.095	0.391

Note. t-test: TMT-A, Trial Making test A; TFT, Tangle Figure Test.

Figure 1

Description of the full RCT process.



All participants were taking medication during the intervention period (Table 3). At baseline, the mean number (SD) of medications per resident was 5.81 (2.62), respectively 5.56 (2.5) for the MT group and 6.11 (2.81) for the ST group.

Table 3

Medication exposure of the study population during the intervention period (5-month)

Type of Drugs	Music Therapy group (N = 23)		Story Telling group (N = 19)	
	N(%)	Mean (SD)	N(%)	Mean (SD)

Cardiovascular: Angiotensin II Receptor Blockers (ARBs) Angiotensin-Converting Enzyme (ACE) Inhibitors Calcium Channel Blockers Diuretics Cardiac Glycosides Anticoagulants	17 (73.9)	.73(.44)	15 (78.9)	.79 (.41)
Bone and Mineral Metabolism: Bisphosphonates Vitamin D Mineralocorticoids	16 (69.6)	.69(.47)	13 (68.4)	.68(.47)
Central and Autonomic Nervous System: Antidepressants Antipsychotics Hypnotics Alpha Blockers Antiepileptics Cholinesterase Inhibitors Dopamine Agonists	9 (39.1)	.39 (.49)	8 (42.1)	.42 (.51)
Pain Management: Antiplatelets Analgesics Opioid Agonists NMDA Receptor Antagonists	16 (69.6)	.69 (.47)	15 (78.9)	.78 (.41)
Lipid Modifiers: Statins	8 (34.8)	.34 (.48)	6 (31.6)	.31 (.47)
Allergy and Respiratory Agents: Antihistamines	5 (21.7)	.21 (.42)	4 (21.1)	.21 (.41)
Hormones and Modulators: Antiandrogens Thyroid Hormones	11 (47.8)	.47 (.51)	10 (52.6)	.52 (.51)
Gastrointestinal Agents: Prostaglandin Analogs Antidiarrheals Stimulant Laxatives H2 Antagonists	17 (73.9)	.73 (.44)	14 (73.7)	.73 (.45)
Antidiabetics	6 (26.1)	.26 (.44)	3 (15.8)	.15 (.37)
Antimicrobials	2 (8.7)	.98 (.28)	1 (5.3)	.05 (.22)

Interventions

Improvisational music therapy (MT) and matched storytelling (ST) protocol was developed (more details below). All the care homes involved in the project followed the same protocol to minimize random noise effects in the data. Both music therapists and storytellers were part of the staff of the care home provider and underwent training on their respective protocols to guarantee consistency in the project and protocol's fidelity. In-person training, led by the first author in two sessions, allowed trainees to familiarize themselves with the material and seek

clarification. The researcher remained available for consultation throughout the intervention, and online email forums were established for peer support, separately for music therapists and storytellers.

The experimental group attended sixteen one-to-one active 45-minute improvisational MT sessions, typically held once a week over a five-month period to accommodate the therapists' annual leave. Four qualified music therapists, registered with the UK Health and Care Professions Council (HCPC), delivered the sessions. There are different MT approaches (for a review, see Darrow, 2008). For this study, we employed an interactive MT approach with improvisational techniques (Hsu et al., 2015) that is used by the care provider's MT services. These consist of interactive music activities, where the therapist actively engages with the participants by singing, playing musical instruments and asking music-related or personal questions that help the participant to reflect on aspects of the music and their own life. In the care provider's MT service, improvisation is an integrative activity of the musical practice. Musical improvisation is the process during which rhythms and melodies hitherto unknown to both the improviser and their listeners, are produced for the first time. The process behind the improvisation and the product (the improvisation itself and what is influenced by it) is closely related to the subject that improvises, who expresses personal and listener's feelings through sounds and silences in relation to the environment (Biasutti & Frezza, 2009). Musical improvisation helps players to develop abilities such as attention, autobiographical memory, working memory, language and communication, discriminative abilities, divergent thinking, cognitive flexibility and motor skills (Beaty, 2015; Biasutti, 2017; Biasutti & Frezza, 2009; Diaz Abrahan et al., 2020; Van Puyvelde et al., 2014b). In MT, improvisation is used to develop social skills, intrinsic motivation, cooperative learning and music abilities, as well as concentration, creativity process, self-consciousness and self-control (Bugos et al., 2007; McPherson, 2005; Kim, Wigram & Gold, 2008; Rodrigues et al., 2013; Roden et al., 2014). Thus, it may provide benefits to mood and psychiatric symptoms (Aalberset al., 2019; Hsu, et al., 2015). In addition to improvisation, the MT approach is grounded in an affect-matching strategy (Franco et al., 2014; Swaine, 2014; Ward et al., 2021). Improvements in cognitive performance will be facilitated by music that is "perceived by the listener to express emotions that are congruent rather than incongruent with the listener's actual affective state" (Franco et al., 2014; p. 872). Within the context of improvisational MT practice, the therapist creates music designed to harmonize with the client's inner state (Ansdell & Pavlicevic, 2005). This occurs within an environment of trust, specifically tailored to address the client's needs (Wigram, 2004)."

Following Chanda and Levitin (2013) and Moreno-Morales's (et al., 2020) recommendation, the control group had a parallel non-music experimental intervention designed to match attentional engagement, intensity level, and mood state modification to that of the MT intervention. Therefore, the control group participated in sixteen 45-minute story-telling sessions, typically held once a week over a five-month period to accommodate the therapists' annual leave. The ST protocol was developed to align with the improvisational and affect-matching approach employed in the MT sessions, ensuring comparable levels of participant engagement and emotional resonance between the two interventions.

Four expert activity coordinators with degrees in psychology or occupational therapy conducted the sessions. Similar to the MT sessions, the ST activities included an interactive reading component, where the reader actively engaged with the listener by creating different stories, asking story-related or personal questions that facilitated reflection on aspects of the story and the listener's own life — mirroring the approach used in MT sessions.

Both MT and ST activities aimed to provide participants with an individual moment of well-being where they could share their feelings related to the stories and their personal life-histories.

To maintain the therapeutic relationship, both MT and ST participants received sessions from the same therapist/story tellers throughout the project (Hsu et al., 2015). Activities were conducted in a comfortable and bright room, a relatively quiet context and free of potentially interfering or disturbing stimuli (see *Supplementary Materials* for protocol details).

Both groups continued to receive the standard care offered by the care homes, including daily activities carried out by professional activity coordinators/caregivers. Standard care activities generally included personal care, puzzles games, arts and crafts, crosswords, gentle physical games, group live music listening (once a week), mass service, and Bible reading. No ST activities or MT sessions have been offered to the participants outside the intervention session. After the intervention period, a course of the MT treatment was offered to the participants in the control group.

This structured approach aimed to control for variables that could influence outcomes and ensure comparability between MT and ST interventions, facilitating a more rigorous evaluation of the specific effects of music training on cognitive and emotional functions.

Outcome Measures

A neuropsychological test battery was administered to the participants at baseline and one week after the ending of the intervention period. The tests were administered by three psychologists, experts in cognitive testing with older adults, who were masked to treatment conditions.

Psychological tests

Screening tests, used for recruitment.

Three screening tests were used in the pre-test phase to balance groups for physical pathological level, cognitive reserve and general cognitive functions.

CIRS, Cumulative Illness Rating Scale (Linn et al., 1968) assesses health/physical illness severity in older adults. Across 14 categories, it considers medical history and symptoms, yielding a cumulative score (0-56). A higher score signifies greater symptom severity. Demonstrating good reliability, α ranges from 0.55 to 0.91 (Linn et al., 1968).

CRIq, Cognitive Reserve Index Questionnaire (Nucci et al., 2012), quantifies cognitive reserve through a standardized interview with the participant or caregiver. Gathering demographic details about education, work, and free-time activities, it combines these into a single index (mean = 100, SD = 15). Employed in the pre-test to ensure balanced groups, it demonstrates good reliability with α = 0.62 (Nucci et al., 2012).

MMSE, Mini-Mental State Examination (Folstein et al., 1975) a standardized test to measure general cognitive functions. The total score is between 0 to 30 points with scores from 26 indicating cognitive normality. MMSE has a good internal consistency with α between 0.6 to 0.9 (Mitchell, 2013).

Pre-post cognitive tests

The primary pre/post cognitive test was the MCT, Music Cognitive Test (Mangiacotti et al., 2022). This test is designed to measure potential changes resulting from music-based interventions and can also effectively screen cognitive functions. It exhibits a robust correlation with widely used cognitive assessment tools such as MMSE, MoCA, and the Severe MMSE (Harrell et al., 2000). This test demonstrates higher sensitivity in detecting cognitive changes compared to MMSE, attributed to the presence of direct subscales measuring executive functions abilities, such as divided and alternating attention and inhibitory processes. According to its validation results the MCT has shown a robust correlation with both the MoCA (r = .785, p < .001) and MMSE (r = .810, p < .001) scores, displaying an excellent sensitivity in identifying cognitive impaired individuals according to both MMSE and MoCA diagnostic

criteria (99.4% and 93.0%, respectively) and presents a good internal consistency with α = .895 (Mangiacotti et al., 2022). The total score is between 0 to 52 points with scores from 45 indicating cognitive normality.

Several secondary measures were used for convergent evidence and to explore specific areas of cognition:

MoCA, Montreal Cognitive Assessment (Nasreddine et al., 2005), is an extensively validated screening tool for neuro-cognitive impairment, widely used in clinical and experimental settings. Assessing executive function and visuospatial abilities, the scoring ranges from 0 to 30 points, with 26 and above indicating cognitive normality. It demonstrates good internal consistency with α = .903 (Freitas et al., 2012).

VFT, Verbal Fluency test, Phonemic & Semantic (Ardila et al., 2006), assesses lexical skills, semantic verbal fluency, and strategic organization. In 60 seconds, participants generate words starting with a designated letter. The total score is the average words per letter (three in total), excluding repetitions. The test demonstrates good internal consistency with Cronbach's alpha .86 (Aki et al., 2022).

CDT, Clock Drawing test (Mondini et al., 2011, adaptation of Critchley, 1953), to evaluate praxis abilities, mental representation and planning skills. In this version, the total score ranges between 0 (minimum) to 10 points (maximum). Internal consistency α = .93 (Mondini et al., 2011).

TFT, Tangled Figure Test (in Mondini et al., 2011, adaptation of Rey, 1958), provides information on spatial-cognitive abilities and executive and naming difficulties. The participant has 4 minutes to discriminate as many pictures as possible from a tangled picture. The total number of pictures is 50. The total score is given by the number of pictures correctly identified and named. The test presents a good internal consistency with α = .89 (Mondini et al., 2011).

TMT-A, Trail Making Test (in Mondini et al., 2011, adaptation of Spreen & Strauss, 1998) to assess selective attention and psychomotor speed. The total score is based on the number of seconds needed to complete the test. Internal consistency α = .96 (Mondini et al., 2011).

Pre-post Mood and Behavioural Tests

The primary pre/post behavioural test was the CSDD - Cornell Scale for Depression in Dementia (Alexopoulos et al., 1988), to screen depressive symptoms in older adults with dementia and cognitive impairment. Administered to both participants and care staff, the 19-item scale uses a 3-point rating (absent, mild or intermittent, severe). With a total score derived from the sum of each item, a cut-off of 10 indicates a probable major depressive episode, while

18 indicates a definite major depressive episode. The test exhibits good internal consistency with α = 0.84 (Williams & Marsh, 2008).

BADL, Bristol Activities of Daily Living Scale (Bucks et al., 1996), developed for evaluating activities of daily living (ADL) through care staff interviews. Tailored for individuals with mild dementia in care home settings, the 20-item questionnaire appraises basic ADL such as food preparation, eating, drink preparation, mobility, and orientation to time. Using a 4-point severity scale (0 to 3) for each category, the total score ranges from 0 to 60, demonstrating internal consistency between 0.901 and 0.940 (Boyd et al., 2018).

QOL, Quality of Life in Alzheimer's Disease (Logsdon et al., 2002), assesses the quality of life in dementia through participant and caregiver reports. Comprising 13 items rated on a 4-point scale (poor – excellent), total scores range from 13 to 52. The test provides separate participant and caregiver scores, with an option to calculate a composite score. Demonstrating reliability and validity for individuals with MMSE greater than 10 (mild to moderate cognitive impairment), the QOL-AD exhibits good internal consistency with α = 0.82 (Buasi & Permsuwan, 2014).

SWLS, The Satisfaction With Life Scale (Diener et al., 1985), is a widely employed self-report measure for global life satisfaction, demonstrating sensitivity to changes during clinical interventions (Pavot & Diener, 2009). Comprising 5 items rated on a 7-point agreement scale (strongly disagree - strongly agree), the scale exhibits good internal consistency with α = 0.74 (López-Ortega et al., 2016).

NPI – The Neuropsychiatric Inventory (Cummings et al., 1994) assesses common behavioural symptoms in individuals with cognitive impairment, providing information directly from caregivers. It demonstrated moderate internal consistency (α between 0.55 and 0.68; Reuther et al., 2016). The NPI yields two indexes: one reflecting the frequency and severity of observed behavioural symptoms in the participant (score range 0-144), and the other gauging the caregiver's perceived disruptiveness (increased stress level or additional workload) in response to observed behaviour (score range 0-60). For both indexes, a higher score indicates a greater severity of symptoms. In addition to the behavioural total score, four subgroups scores can be calculated according to the factor analysis conducted by Aalten (et al., 2007): i) psychotic (NPI domains: hallucinations, delusion), ii) hyperactivity (NPI domains: agitation, euphoria, disinhibition, irritability, aberrant motor behaviour), iii) apathy (NPI domains: apathy, eating abnormalities) and iv) affective (NPI domains: depression, anxiety).

MOSES, Multidimensional Observation Scale for Elderly Subjects (Helmes et al., 1987) was developed to assess behaviour in older adults through an interview with the caregiver. Based on the authors' validation, the test demonstrates good internal consistency in the 0.8 range. It comprises five subscales: self-care, disorientation, depression, irritability and social withdrawal. However, only the last two subscales were considered in this study since the first domains were already evaluated using the BADL and NPI scales. Each subscale consisted of 7 or 8 questions, rated on a scale of 1-5 (not observed - always observed behaviour).

Biomarkers

Saliva samples and cardio-respiratory measures were collected from participants to investigate possible endocrinological and physiological effects of the interventions.

Cortisol/DHEA analyses

Daily saliva samples were collected at three different time points during the 4-month intervention period: pre-intervention (T0), after two months (T1) and post-intervention (T2). Due to the cortisol and DHEA dependency on circadian rhythm, on each collection day, four saliva samples were collected: i) within 10min from participants waking up (6am - 8am), ii) before lunchtime (11.00am - 12.30am), iii) before dinner (4pm - 5pm), iv) evening (7pm - 8pm). It was possible to collect the first morning sample because we knew the participants' wake-up time. However, due to limited resources (e.g., care home staff/researchers' availability), we were unable to collect the awakening response (30 minutes after waking up). Saliva samples were collected through passive drool in Eppendorf 2ml tubes. Following Bhattarai et al. (2018), we implemented measures to ensure a clean saliva sample. Participants refrained from eating or brushing their teeth 60 minutes before collection, with no drinking at least 30 minutes prior. To minimize mucous and remove food particles, participants rinsed their mouths with water at least 10 minutes before collection. This process helped balance saliva in the mouth, preventing overly diluted samples. Collected samples were initially stored at 4°C, then frozen at -80°C within 16 hours for subsequent shipment and examination. Analyses of the cortisol and DHEA were conducted in collaboration with Dresden Lab GmbH service (Germany), and samples were analysed using high-performance liquid chromatography-tandem mass spectrometry (LC-MS/MS, Keevil, 2013) with atmospheric pressure chemical ionization coupled with online solid-phase extraction.

ECG and respiration

Three-minute resting-state ECG (electrocardiogram) and respiration measurements were collected before and after the 4th and 15th MT or ST sessions to understand whether the activities could lead to short-term changes in the parasympathetic activity. In order not to interfere with the development of the therapeutic alliance between the therapist and participant, the pre-post session change was measured at the 4th session instead of the 1st (as recommended, e.g., in Hsu et al., 2015). Since the end of the treatment course is considered a very meaningful moment, with a strong symbolic value, we decided to collect ECG and respiration measures on the 15th session rather than the last one for not impacting participants' therapeutic experience.

ECG and respiration measures were obtained using the portable BioRadio™ system (Cleveland Medical Devices, Inc., OH), a wearable biomedical device for recording physiological signals. RSA (Respiratory Sinus Arrythmia) values were analysed using VivoSense software (Vivonoetics, San Diego, USA). Detailed information regarding the collection, filtering, and data analysis of the ECG-respiratory signal is provided in the *Supplementary Material*.

Results

Data of the neuropsychological, well-being and biomarker measures were analysed using a mixed design ANOVA with time (pre-post intervention) as within-subject factor and group (MT vs ST) as between-subjects factor. The dependent variables were the cognitive, neuropsychological and biomarkers tests. Quantitative data were processed using IBM SPSS Statistic 25. Partial eta-square (η_p^2) and Cohen's *d* were used as a measure of effect size. All data were tested on normality (Kolmogorov–Smirnoff test), homogeneity and sphericity. Results of the pairwise Bonferroni post-hoc comparison for all tests are reported in Table 4. Recommendations from Tian et al. (2018) have been considered for post-hoc comparison when omnibus tests were not significant. Dynamic Bayesian Network has been considered to unravel the probabilistic dependencies among variables over time. The SPSS and R code used to analyse the findings can be found at: https://doi.org/10.22023/ANONYMISED

Table 4

Bonferroni Post-hoc analyses of the Mixed Method ANOVA on the cognitive and behavioural tests

	Music Therapy group						Story Telling group					
Measure	**n**	**Baseline** mean(SD)	**Post** mean(SD)	***p*(η_p^2)**	**MD** (pre-post)	**95% CI**	**n**	**Baseline** mean(SD)	**Post** mean(SD)	***p*(η_p^2)**	**MD**	**95% CI**
Cognitive												
MCT	23	32.52(6.90)	36.65(4.09)	.0001(0.371)*	-4.13	-5.84 to -2.41	19	34.53(4.55)	34.68(5.27)	.867(0.001)	-.158	-2.04 to 1.73
MOCA	23	10.39(3.77)	12.35(3.96)	.001(0.234)*	-1.96	-3.08 to -.83	19	12.89(5.09)	12.74(3.95)	.799(0.002)	.158	-1.08 to 1.40
VFP	23	7.23(3.67)	9.32(3.71)	.0001(0.365)*	-2.087	-2.97 to -1.20	19	7.89(4.81)	9.18(4.92)	.010(0.153)*	-1.289	-2.25 to -.321
VFS	23	7.65(3.52)	8.00(3.21)	.590(0.007)	-.348	-1.64 to .94	19	7.68(4.55)	7.89(5.91)	.766(0.002)	-.211	-1.63 to 1.21
CDT	23	4.82(3.22)	6.04(2.86)	.037(0.104)*	-1.217	-2.36 to -.07	19	5.61(3.49)	6.26(3.41)	.296(0.027)	-.658	-1.91 to .59
TFT	19	8.05(5.15)	10.37(5.51)	.036(0.138)*	-2.316	-4.47 to -.15	13	12.46(7.96)	12.54(6.54)	.952(0.0001)	-.077	-2.68 to 2.25
TMT-A	18	187.61(149.21)	150.06(109.23)	.043(0.133)*	37.550	1.172 to 73.928	13	151.11(112.22)	152.51(83.03)	.947(0.0001)	-1.398	-44.20 to 41.20
Behavioural and Well-being												
CSDD	23	7.17(4.30)	3.22(2.17)	.0001(0.324)*	3.957	2.13 to 5.78	19	9.47(5.26)	8.89(4.58)	.563(0.008)	.579	-1.43 to 2.59
BADL	23	14.00(5.34)	14.43(6.11)	.701(0.004)	-.435	-2.71 to 1.83	19	15.21(8.46)	14.74(7.80)	.704(0.004)	.474	-2.02 to 2.97
QoL	23	34.46(5.09)	36.40(5.51)	.048(0.094)*	-1.942	-3.86 to -.02	19	32.92(5.18)	34.05(4.48)	.290(0.028)	-1.123	-3.24 to .99
SWLS	23	24.17(6.11)	24.52(7.97)	.795(0.002)	-.348	-3.04 to 2.34	19	21.26(6.25)	21.74(6.76)	.748(0.003)	-.474	-3.43 to 2.48
NPI	23	12.65(16.83)	8.43(9.75)	.268(.031)	4.217	-3.38 to 11.81	19	11.58(9.38)	18.95(19.65)	.082 (0.074)	-7.368	-15.72 to .98
NPId	23	4.04 (5.41)	3.04(4.48)	.436(0.015)	1.00	-1.57 to 3.57	19	4.68(4.83)	5.37(6.13)	.627(0.006)	-.684	-3.51 to 2.14
NPIh	23	6.21(6.64)	4.00(6.46)	.211(0.039)	2.217	-1.31 to 5.74	19	6.26(7.24)	7.57(10.92)	.497(.012)	-1.316	-5.19 to 2.56
NPIp	23	3.73(6.24)	2.00(3.42)	.194(0.042)	1.739	-.92 to 4.40	19	2.68(3.83)	2.63(5.58)	.971(.0001)	.053	-2.87 to 2.98
NPIa	23	2.86(6.26)	1.04(1.63)	.106(0.064)	1.826	-.40 to 4.05	19	2.57(3.25)	3.68(4.94)	.388(0.020)	-1.105	-3.55 to 1.34
NPIap	23	3.13(5.57)	1.39(2.31)	.215(0.038)	1.739	-1.05 to 4.53	19	2.63(3.26)	5.05(6.76)	.119(0.060)	-2.421	-5.49 to .64
MOSES IV	23	11.52 (3.75)	11.35(4.52)	.899 (0.001)	.174	-2.58 to 2.92	19	11.95 (4.68)	17.05(8.01)	.002(0.225)*	-5.105	-8.13 to -2.08
MOSES V	23	15.57(4.63)	15.74(4.48)	.353(0.027)	-.174	-2.54 to 2.20	19	17.05(4.10)	17.79(5.31)	.132(0.069)	-.737	-3.35 to 1.87

Note. SD, Standard deviation; MD, Mean difference; CI, Confidence intervals; MCT, Music Cognitive Test; MOCA, Montreal Cognitive Assessment; VFP, Verbal fluency phonetic; VFS: Verbal fluency semantic; CDT, Clock Drawing Test; TFT, Tangle Figure Test; TMT-A, Trial Making Test-A; CSDD, Cornell Scale for Depression in Dementia; BADLS, Bristol Activities of Daily Living Scale; QoL, Quality of Life in Alzheimer's Disease; SWLS, Satisfaction With Life Scale; NPI, Neuropsychiatric inventory; NPId, Neuropsychiatric inventory disruptiveness; NPIh, Hyperactivity; NPIp, Psychosis; NPIa, Affective; NPIap, Apathy; MOSES IV, Multidimensional Observation Scale for Elderly Subjects – Irritability; MOSES V, Multidimensional Observation Scale for Elderly Subjects - Social Withdrawal.

*p<0.05; **p<0.001.

Neuropsychological outcomes

Results of the main analyses showed a general improvement (p<0.05) with medium to large effect size for the experimental group (MT) in almost all tests compared to the control group (ST) (p>0.05) (Table 5).

Table 5

Mixed design ANOVA results of cognitive test changes. Within-subjects factor: time (pre-post intervention); between-subjects factor: group (MT or ST).

Dependent Variable	Source of Variance	SS	d.f.	F	*p*	η_p^2
MCT	Time	95.671	1	11.507	.002*	0.223
	Group	.007	1	0.200	.991	0.000
	TimexGroup	82.099	1	11.507	.003*	0.198
	Error	332.568	40			
MOCA	Time	16.830	1	4.683	.036*	0.105
	Group	43.524	1	1.377	.248	0.033
	TimexGroup	23.259	1	6.472	.015*	0.139
	Error	143.741	40			
VFP	Time	59.309	1	27.193	.001*	0.405
	Group	1.373	1	0.040	.842	0.001
	TimexGroup	3.309	1	1.517	.225	0.037
	Error	87.242	40			
VFS	Time	1.622	1	0.345	.560	0.009
	Group	0.028	1	0.001	.977	0.001
	TimexGroup	0.098	1	0.021	.886	0.001
	Error	188.188	40			
CDT	Time	18.295	1	4.979	.031*	0.111
	Group	5.190	1	0.229	.062	0.007
	TimexGroup	1.629	1	0.443	.509	0.011
	Error	146.970	40			
TFT	Time	22.095	1	2.081	.159	0.065
	Group	167.044	1	2.529	.122	0.078
	TimexGroup	19.345	1	1.822	.187	0.057
	Error	318.514	30			
TMT-A	Time	4932.831	1	1.732	.198	0.056
	Group	4376.359	1	0.173	.680	0.006
	TimexGroup	5725.160	1	2.011	.167	0.065
	Error	82570.55	29			

Note. MCT, Music Cognitive Test; MOCA, Montreal Cognitive Assessment; VFP, Verbal fluency phonetic; VFS: Verbal fluency semantic; CDT, Clock Drawing Test; TFT, Tangle Figure Test; TMT-A, Trial Making Test-A.

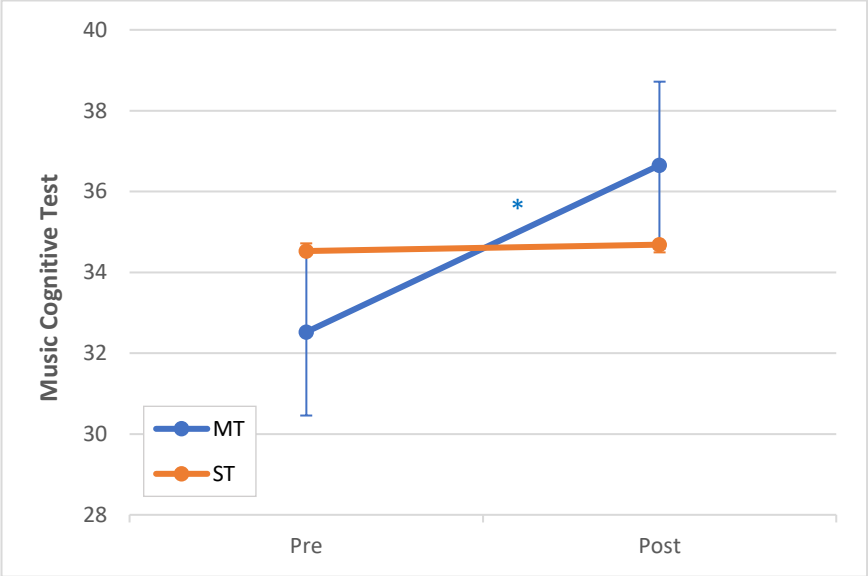
** p*<0.05.

Significant differences between MT and ST were found in the primary measure, MCT (see Figure 2), as well as converging results for MoCA and CDT: improved cognitive scores were recorded post-intervention in the MT but not in the ST group. Concerning Verbal Fluency, both groups presented a significant pre/post difference in

performance. Post-hoc pairwise Bonferroni comparisons suggested a higher pre/post improvement for the MT (p<.001) compared to the ST group (p<.05). No changes were observed in the Semantic Verbal Fluency task. Regarding the TFT (spatial, executive and naming abilities) and the TMT-A (selective attention and psychomotor speed), despite the main effects not being fully significant (p>.05), pairwise Bonferroni comparison analyses indicated a significant improvement in the performance over time for the MT group (p<.05) while the ST group remained stable (p>.05) (Table 4).

Figure 2

MCT scores at pre-post intervention for the MT and ST groups.



Behavioural and Wellbeing Outcomes

Results of the main analyses (Table 6) showed a general improvement (p<0.05) for the MT group compared to the ST group in almost all behavioural and wellbeing tests. Significant differences between MT and ST were found for CSDD, QoL, NPI and MOSES, while no differences were found in NPI-Disruptiveness, SWLS and BADL for both the groups.

Table 6

Mixed design ANOVA results of behavioural test changes. Within-group factor: time (pre-post intervention); between-group factor: group (MT or ST).

Dependent Variable	Source of Variance	SS	d.f.	F	*P*	η_p^2
CSDD	Time	107.015	1	11.421	.002*	0.222
	Group	331.050	1	13.073	.001*	0.246
	TimexGroup	59.349	1	6.334	.016*	0.137

	Error	374.794	40			
QoL	Time	48.867	1	4.691	.036*	0.105
	Group	78.606	1	1.882	.178	.045
	TimexGroup	3.491	1	.335	.566	0.008
	Error	416.651	40			
SWLS	Time	3.511	1	0.172	.680	0.004
	Group	168.768	1	2.302	.137	0.054
	TimexGroup	0.082	1	0.004	.950	<0.001
	Error	815.977	40			
BADL	Time	0.008	1	0.001	.982	<0.001
	Group	11.903	1	0.146	.704	0.004
	TimexGroup	4.294	1	0.296	.590	0.007
	Error	581.195	40			
NPI	Time	51.654	1	0.318	.576	0.008
	Group	463.540	1	1.788	.189	0.043
	TimexGroup	698.321	1	4.301	.045*	0.097
	Error	6494.167	40			
NPI - Disruptiveness	Time	0.519	1	0.028	.868	0.001
	Group	45.756	1	1.273	.266	0.031
	TimexGroup	14.757	1	0.794	.378	0.019
	Error	743.053	40			
NPI - Hyperactivity	Time	4.229	1	0.121	.730	0.003
	Group	68.352	1	0.764	.387	0.019
	TimexGroup	64.943	1	1.853	.181	0.044
	Error		40			
NPI - Psychosis	Time	16.702	1	0.836	.366	0.020
	Group	0.932	1	0.033	.858	0.001
	TimexGroup	14.797	1	0.741	.394	0.18
	Error	798.691	40			
NPI - Affective	Time	2.703	1	0.193	.663	0.005
	Group	28.733	1	1.155	.289	0.028
	TimexGroup	44.703	1	3.196	.081	0.074
	Error	559.547	40			
NPI - Apathy	Time	2.419	1	0.110	.741	0.003
	Group	52.030	1	2.207	.145	0.052
	TimexGroup	90.038	1	4.109	.049*	0.093
	Error	876.533	40			
MOSES - irritability	Time	126.513	1	5.936	.019*	0.129
	Group	195.517	1	5.363	.026*	0.118
	TimexGroup	144.989	1	6.803	.013*	0.145
	Error	852.547	40			
MOSES – social withdrawal	Time	4.315	1	0.272	.605	0.007
	Group	65.112	1	2.392	.130	0.056
	TimexGroup	1.649	1	0.104	.749	0.003
	Error	635.494	40			

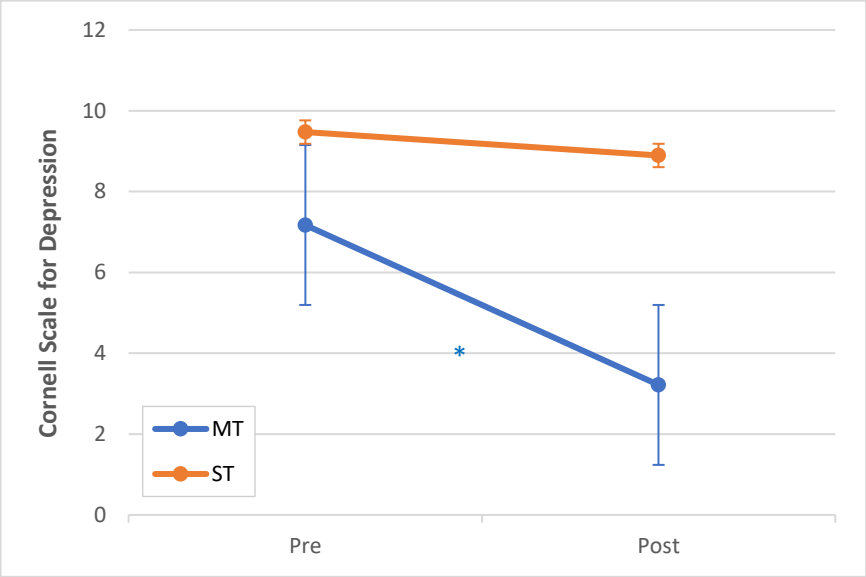
Note. QoL, Quality of Life in Alzheimer's Disease; SWLS, Satisfaction With Life Scale; BADLS, Bristol Activities of Daily Living Scale; CSDD, Cornell Scale for Depression in Dementia; NPI, Neuropsychiatric inventory; MOSES, Multidimensional Observation Scale for Elderly.

*p<.05.

Results on CSDD showed a reduction in the score over time for the MT group compared to the ST group that instead remained stable, indicating an improvement in depressed mood symptoms for the MT group (Figure 3). This was reflected in the statistical difference across groups (MT/ST) after the end of the intervention period with a large effect size (-2.3, 95% CI [-7.85 to -3.51], $p < .001$, $\eta_p^2 = 0.410$).

Figure 3

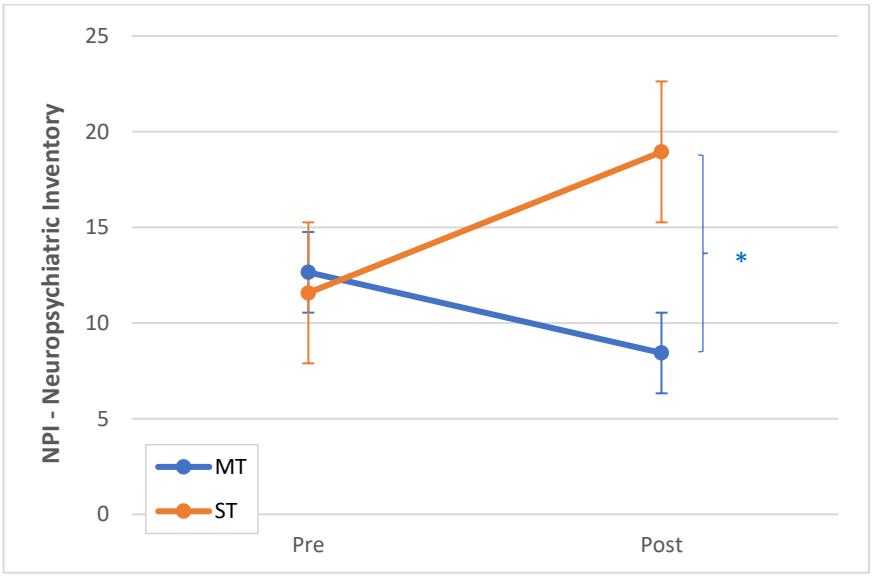
Cornell Scale for Depression in Dementia (CSSD) data at pre-post intervention for Music Therapy (MT) and Story Telling (ST) groups.



Regarding the NPI, a mean score difference is notable (Figure 4), with an improvement in the experimental group and aggravation of symptoms for the control group along the intervention period. Despite pairwise Bonferroni post-hoc comparisons were not significant for either the MT or ST group over time, the difference across the experimental and control groups after the intervention period was significant with a medium effect size (-10.513, 95% CI [-19.93 to -1.09], p = .030, $\eta_p^2 = 0.113$). When considering the four NPI sub-scores separately (Hyperactivity, Psychosis, Affective and Apathy), an improvement in the means of the experimental group is noted for the *Affective* and *Apathy* sub-scores. In particular, the *time*X*group* interaction approached significance for the sub-score *Affective* (p=.084) and was fully significant for *Apathy* (p<.05). Furthermore, a significant difference between groups at the end of the intervention emerged for sub-scores *Affective* (-2.641, 95% CI [-4.85 to -.42], p = .021, $\eta_p^2 = 0.127$) and *Apathy* (-3.661, 95% CI [-6.71 to -.62], p = .020, $\eta_p^2 = 0.129$). Scores related to caregiver perceived disruptiveness showed no significant changes (see Table 4 & 6).

Figure 4

Overall NPI scores at pre-post intervention for MT and ST groups.



Scores for the MOSES subscales *irritability* and *social withdrawal* were analysed separately. Results on the sub-scale *irritability* showed a significant main effect of *time*, *group,* and in the interaction between the two factors ($p<.05$). In the context of pairwise Bonferroni post-hoc comparisons, the results indicated an increase in irritability scores (participants feeling more irritated) within the control group. ($p<.05$), while the experimental group remained stable. Lastly, the difference across groups at the end of the intervention was significant with a medium effect size (-5.705, 95% CI [-9.67 to -1.73], p =.006, η_p^2 = 0.174). Conversely, results on the subscale *social withdrawal* showed only a tendency to the significance for the main factor of *time* and no difference for *group*, between the two factors and in the Pairwise Bonferroni post-hoc comparisons analyses, indicating a possible influence of the MT in improving people engagement in social activities.

Dynamic Bayesian Network

The aim of this analysis was to unravel the probabilistic dependencies among variables over time. A Dynamic Bayesian Network analysis (DBN) has been implied. (Information about the DBN analysis is reported in the *Supplementary Material*).

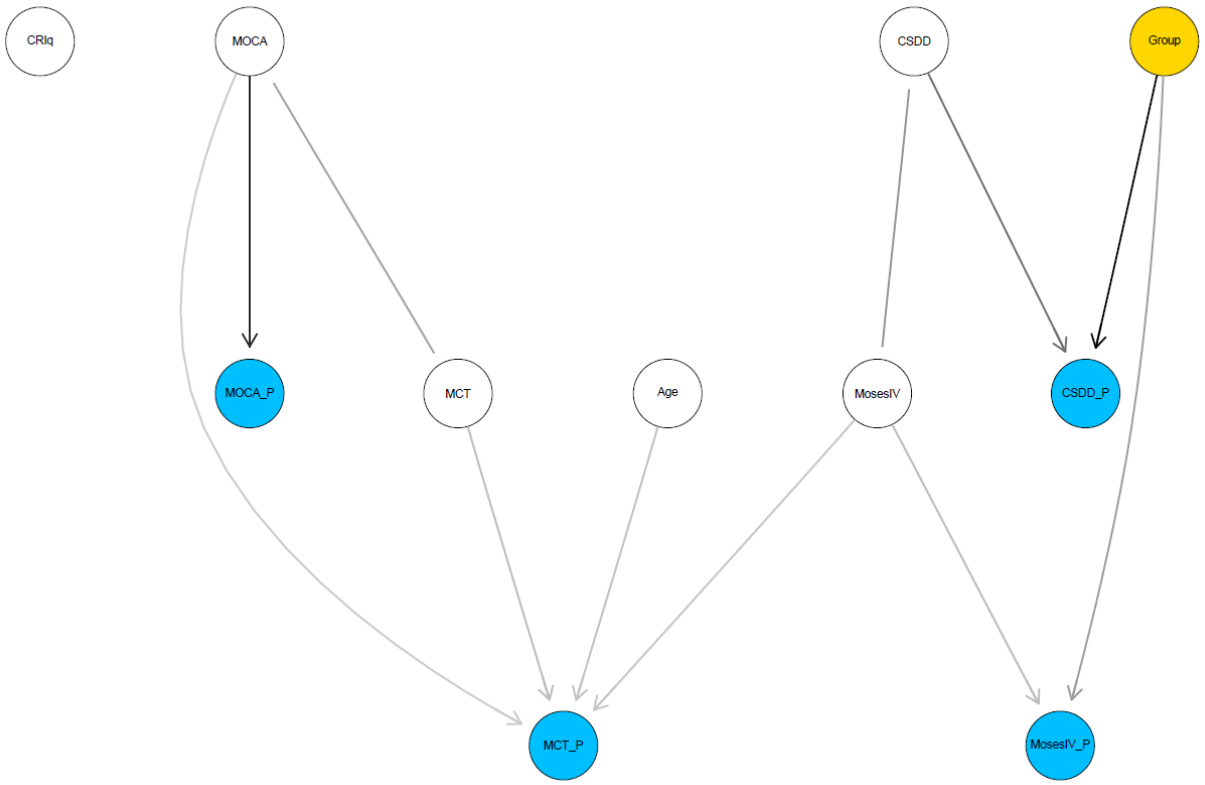
For this analysis, six primary variables were selected: the between-subject variable group, along with two main demographic variables (age and CRIq), and four cognitive-behavioral variables (MCT, MOCA, CSDD, MOSES-IV). These variables were chosen based on their significance in the ANOVA for both *Time* and *TimexGroup* factors. The dataset underwent quantization by applying thresholds to these variables, categorizing subjects based on

values falling below or above specified thresholds. The thresholds include: group (MT, ST); Age = 80 (per Chu et al., 2013), CRIQ = 100 (standardized mean for CRIq), MCT = 33 (cut-off moderate cognitive impairment), MOCA = 10 (cut-off moderate cognitive impairment), CSDD = 10 (depressive episode cut-off), and MOSES-IV = 11 (total baseline sample mean). The MOSES-IV threshold was set at the average of the total sample, given the absence of a reported cut-off.

The analysis unveiled key direct dependencies, with the DBN indicating strong probabilistic links between group and post-intervention CSDD scores, as well as baseline MOCA values and post-intervention MOCA values (Figure 5). Group also demonstrates a notable dependency on post-intervention MOSES-IV. Over time, dependencies emerge between pre/post-intervention CSDD values, pre/post-intervention MCT values, and baseline MOCA values on post-intervention MCT values. Additionally, Age directly influences the post-intervention MCT score. Notably, no direct dependencies were found for CRIq. The analysis also identified indirect dependencies, including MoCA baseline - MCT baseline - MCT post-intervention and CSDD baseline – MOSES-IV baseline – MCT post-intervention scores.

Figure 5

Dynamic Bayesian Network describing the dependencies among variables over time.



Note. In blue, the variables of interest at the end of the intervention period. The edges thickness is proportional to the probabilistic dependency strength. CRIq, Cognitive Reserve Index Questionnaire; MOCA, Montreal Cognitive Assessment pre-intervention score; MOCA_P, Montreal Cognitive Assessment post-intervention score; CSDD, Cornell Scale for Depression in Dementia pre-intervention score; CSDD_P, Cornell Scale for Depression in Dementia post-intervention score; MCT, Music Cognitive Test pre-intervention score; MCT_P, Music Cognitive Test post-intervention score.

Biomarkers

Hormones

Salivary data were natural-logarithmically transformed to correct for positive skew and to enable general linear model-based analyses. Missing data were estimated using SPSS multiple imputations (Regression Method). Cortisol/DHEA average ratio was calculated for the 4 daily samples collected at T0 (pre-intervention), T1 (mid-way, 2 months), and T2 (post-intervention). Due to medical problems, three individuals (n=1 MT group, n=2 ST group) did not complete two of the daily saliva collections; therefore, their data were excluded from the analysis. The results for the final sample are reported in Table 7.

A mixed-factor ANOVA was used to interpret MT/ST differences and within-group differences over time (T0-T1-T2).

Table 7

Cortisol/DHEA molar ratio of the four daily samples at T0, T1, T2.

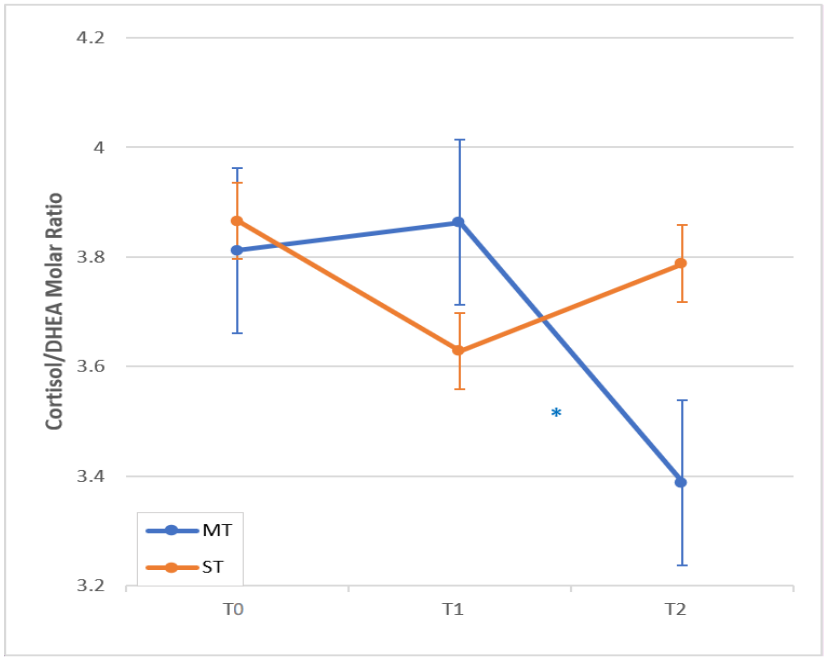
	Music Therapy group (n=22)	**Story Telling group (n=17)**	***p***	**ηp²**
T0 [mean nmol/l (SD)]	3.812 (0.86)	3.866 (1.24)	.875	0.001
T1 [mean nmol/l (SD)]	3.863 (0.75)	3.628 (1.15)	.447	0.016
T2 [mean nmol/l (SD)]	3.388 (0.80)	3.788 (0.94)	.162	0.052

Note. T0 = pre-intervention; T1 = after 2 months (week 8th); T2 = post-intervention (week 16th)

The main effects of *time*, $F(2,74) = 2.022$; $p = .140$; $\eta_p^2 = 0.052$ and *group,* $F(1,37) = 0.070$; $p = .702$; $\eta_p^2 = 0.002$, were not significant. However, the *time* x *group* interaction was significant with a medium effect-size, $F(2,74) = 3.181$; $p = .047$; $\eta_p^2 = 0.079$. A Pairwise Bonferroni post-hoc analyses revealed, for the experimental group, a significant decrease in the ratio with a large effect size between T0/T2 (0.424, 95% CI [.06 to .79], $p = .018$, $d = 0.517$) and T1/T2 (0.475, 95% ci [.03 to .92], $p = .036$, $d = 0.619$) while the ST group remained stable over the time (Figure 5).

Figure 6

Cortisol/DHEA Molar Ratio of the Four Daily Collections for the MT and ST Group.



Note. T0 = pre-intervention; T1 = after 2 months (week 8th); T2 = post-intervention (week 16th)

Respiratory Sinus Arrhythmia (RSA)

Pre-post session cardio-respiratory measures were collected at the 4th and 15th intervention sessions and RSA values have been calculated as illustrated in the Method, with z-transformations used in the analyses. A total of n = 32 participants agreed to take part in the 4th session and n = 34 in the 15th, see RSA z-values in Table 8. Overall, the MT group clearly optimised RSA by the end of the 4th session. At the 15th, the MT group still showed an improvement post-session, whereas the ST group did not show any noticeable changes in RSA at either session.

Table 8

Pre-post session RSA Z-values at the 4th and 15th intervention.

	N	Music Therapy group	N	Story Telling group	*p*	ηp²
RSA – 4th pre[mean(SD)]	16	-0.216(0.68)	16	-0.022(0.91)	.502	0.015
RSA – 4th post [mean(SD)]	16	0.394(1.08)	16	0.022(1.07)	.338	0.031
RSA – 15th pre [mean(SD)]	21	-0.134(0.93)	13	0.012(1.08)	.731	0.004
RSA – 15th post [mean(SD)]	21	0.134(1.04)	13	0.013(0.91)	.733	0.004

Note. RSA: Respiratory Sinus Arrhythmia

Two separate mixed factor ANOVA analyses were conducted for session 4 and 15, with *time* (pre/post session) as within-subject factor and *group* (MT/ST) as between-subjects factor. The results (Table 9) revealed a

significant interaction, with an improvement in the RSA value for the MT group over the session: pairwise Bonferroni post-hoc analysis showed a significant pre-post increase in the RSA value for the MT group at both the 4th (-0.610, 95% CI [-.96 to -.26], p = .001, η_p^2 = 0.296) and 15th session (-0.268 95% CI [-.47 to -.06], p = .011, η_p^2 = 0.184) whilst the ST group did not change, suggesting a gain only in the MT group.

Table 9

Mixed design ANOVA results of RSA pre-post activity at the 4th and 15th sessions. Within-group factor: time (pre-post activity); between-group factor: group (MT or ST).

Dependent Variable	Source of Variance	SS	d.f.	F	*p*	ηp²
RSA – 4th (MT=16; ST=16)	Time	1.714	1	7.261	.011*	0.195
	Group	0.127	1	0.080	.779	0.003
	TimexGroup	1.283	1	5.439	.027*	0.153
	Error	7.080	30			
RSA – 15th (MT=21; ST=13)	Time	0.523	1	4.989	.033*	0.135
	Group	1.347e-11	1	<0.001	1.00	<0.001
	TimexGroup	0.236	1	2.245	.144	0.066
	Error	3.358	32			

Note. RSA: Respiratory Sinus Arrhythmia

Discussion

In this RCT study, we analysed the effects of 4-month MT vs ST (active control) intervention for people with mild-moderate cognitive impairment living in care homes. The two interventions were based on the *improvisational* and *affect-matching* approaches. As suggested by Chu (et al., 2013) and Moreno-Morales (et al., 2020), we used a quantitative approach to collecting convergent evidence from three sources: cognitive-neuropsychological, behavioural and biomarker tests. Positive outcomes of the music intervention were in line with findings from other studies (see below) and contributed additional evidence, further supporting a robust randomized design. These can be epitomised by two original aspects of the present study; [i] the average age of participants was 86.24 y.o., hence enriching the scant literature on the benefits of MT in adults 80+ (Chu et al., 2013), and [ii] this study is the first to include both RSA and salivary cortisol/DHEA levels as biomarkers in the MT field, providing a more comprehensive understanding of the physiological effects of MT interventions.

Concerning the cognitive-neuropsychological tests (research question one), the findings revealed gains over time for the MT compared to the ST group, with MT presenting improvements in general cognitive functions (MCT

& MoCA), executive functions in particular reasoning, abstraction and cognitive flexibility (CDT, TFT), language (phonological verbal fluency), selective attentional abilities and psychomotor speed performances (TMT-A). Conversely, the ST group showed a significant improvement only in the phonological verbal fluency test, remaining generally stable in all the other tests. Considering that, i) the two intervention activities were matched for key aspects and ii) participant groups were well matched at baseline for social-cognitive level, cognitive reserve, physical, behavioural symptoms and wellbeing measurements, it is reasonable to infer that many of the benefits observed in the MT group may be attributable to the nature of the activities themselves (music vs. storytelling). These results are consistent with other studies reporting the cognitive benefits of MT in older adults with neurocognitive disorders (Bian et al., 2021; Chu et al., 2014; Moreno-Morale; Suzuki et al., 2004). it is suggested that active music experiences can enhance a range of cognitive abilities, such as attention (Biasutti & Mangiacotti, 2018; Kim & Yoo, 2019; McPherson, 2005), language (Moreno et al., 2011; Schön et al., 2004), executive functions such as reasoning, abstraction and cognitive flexibility (Bialystok & DePape, 2009; Bugos et al., 2007; Hegde, 2014; Strong and Midden, 2020; Wan & Schlaug, 2010). Furthermore, music experiences can also drive neuroplasticity (Jünemann, et al., 2022; Wan & Schlaug, 2010).

The ST group's enhanced VFP may be linked to the conversational nature of the intervention, while their stability in other measures over time could result from factors like i) the generally supportive nature of the improvisational and affect-matching intervention, or ii) the inherent social aspects in one-to-one interaction. These results underscore how training activities for older individuals with cognitive decline contribute to maintaining cognitive stability, potentially delaying further decline (Wang et al., 2023b).

In terms of behavioural and well-being measures (research question two), the MT group's pre/post-intervention comparisons showed significant improvement in nearly all tests, especially in quality of life (QoL) and depressive mood state (CSDD). Despite a decrease in CSDD scores, participants' average score was initially outside the clinical cut-off zone (indicating no diagnosis of depression), suggesting the decrease might not indicate a change in depression severity but rather an improvement within a non-depressed range. Differences between MT and ST groups emerged in neuropsychiatric symptoms (NPI) at the intervention's end, with MT participants displaying fewer psychiatric symptoms, particularly in the *Apathy* and *Affective* dimensions. Results from NPI subcategories align with those from the CSDD. Regarding MOSES subscales, the control group's *Irritability* scale worsened, while the MT group remained stable over time. The results of MT group in contrast with the one of the

ST are in line with Suzuki et al.'s (2004) results highlighting how MT may manage irritability symptoms among older adults with cognitive impairments. More broadly, our findings on well-being and behavioural symptoms are aligned with previous research showing positive effects of MT in alleviating depressed mood (Chu et al., 2014; Gallego & Garcia, 2017; Guetin et al., 2009; Wang et al., 2023a) and neuropsychiatric symptoms (Hsu, 2015; Gallego & Garcia 2017) while quality-of-life perception is improved (Cooke et al., 2010; Grocke et al., 2009; Porter et al., 2018) in people with cognitive decline living in care home settings. The results of the ST group could be associated with the 'ageing' effect. With ageing neuropsychiatric symptoms such as irritability may worsening over time. This is noted on the mean of the NPI in which the control group had an increase on the NPI score (worsening of the symptoms) over time.

Other wellbeing scales assessing life satisfaction (SWLS) and activities of daily living (BADL) exhibited stability over time for both MT and ST groups. This observation might stem from the broad nature of the items on these scales, which may not be sufficiently sensitive to capture nuanced changes resulting from the interventions. Alternatively, it could suggest that the MT and ST interventions employed in this study might not be particularly effective in eliciting changes in these specific domains. Further research should delve into and clarify these aspects by utilizing different types of scales or by change the nature of the interventions, such as increasing the number of weekly sessions, to explore whether these modifications result in observable changes.

Building on promising cognitive-behavioural findings, Dynamic Bayesian Network (DBN) analysis further explores relationships between observed changes in cognition and behaviour linked to interventions. The DBN highlights a strong influence of group assignment on post-intervention depression scores (CSDD) and irritability scores (MOSES-IV), suggesting the type of intervention directly impacted depressed mood and irritability levels. This aligns with ANOVA results showing a significant reduction in depression for the MT group, contrasting with stability in the ST group, and a stability in irritability for MT while an increase is observed in the ST group. This suggests MT may directly improve mood levels in individuals with cognitive impairment. Regarding cognitive variables (MoCA and MCT), the DBN underscores the strong influence of baseline MoCA scores on post-intervention MoCA scores and between MCT baseline and post-intervention MCT scores, supporting the main ANOVA. Bonferroni post-hoc analysis confirms MT group improvement over time, while the ST group remains stable, aligning with collective findings suggesting the impact of MT interventions on cognition. The DBN reveals a direct relationship between age and post-intervention MCT scores, hinting at potential age influence on cognitive

performance. Due to the small sample size, further analysis to pinpoint age's impact on cognitive performance across groups was not feasible. In intriguing results, the DBN shows no relationship between cognitive-behavioural test scores and cognitive reserve level (CRIq). However, solely relying on these results, it is inaccurate to conclude improvements occurred regardless of cognitive reserve. Larger studies with extended follow-up periods are needed to understand the complex interplay between MT, cognitive reserve, and cognitive functions in individuals with cognitive decline. The DBN unveils an indirect dependency involving MOCA at baseline, MCT Score at baseline, and post-intervention MCT score, suggesting the cognitive state before intervention indirectly influences post-intervention cognitive outcomes, as measured by the MCT. This pattern is also evident in the indirect dependency between pre-intervention CSDD score, pre-intervention MOSES-IV score, and post-intervention MCT, providing insights into potential interconnections between emotional well-being and cognitive performance in the intervention context. Baseline levels of individuals' emotional states (depression and irritability) and the degree of cognitive decline (mild-moderate) may shape responses to interventions, ultimately impacting cognitive abilities as measured by the MCT. Overall, the DBN analysis further supports positive MT impacts on cognitive and behavioural outcomes observed in the initial analysis, offering deeper insights into potential underlying mechanisms.

Concerning the biomarkers (research questions three/four), MT intervention showed influences on both hormonal and physiological measures. Cortisol/DHEA ratio analyses revealed a significant improvement for the MT group (decreased ratio), compared to the ST group, which did not present pre/post intervention changes. These results are in line with previous studies suggesting a positive impact of music and MT in regulating cortisol levels or cortisol/DHEA ratio in general (Fancourt et al., 2016) and specifically in care home residents (Chu et al., 2014; Suzuki et al., 2004; Takahashi & Matushita, 2006). A possible explanation for the efficacy of musical activities in this domain is that they contribute to lowering depressed mood symptoms (see also Gallego & Garcia, 2017; Takahasi & Matsushita 2006; Wang et al. 2023a), which are related with high salivary cortisol/DHEA ratio (Michael et al., 2000; Tafet et al., 2001). Indeed, in the present study, the MT group improved both in the CSDD score (depression) and the cortisol/DHEA ratio, compared to the ST group, which remained stable over time. Another possible explanation is that positive social interactions may influence DHEA levels (Polenick et al., 2021). MT could have been perceived as a pleasant experience by participants impacting DHEA levels and consequently affecting the overall cortisol/DHEA ratio.

The hormonal results may offer insights into the cognitive benefits observed in the MT group. Existing literature recognizes the impact of music listening and playing on steroid levels associated with spatial perception and cognition (Fukui & Toyoshima, 2008). Cortisol-DHEA ratios, as highlighted in studies (Ji et al., 2021; Jin et al., 2016), are linked to neural plasticity. Elevated cortisol levels can potentially harm neurons, particularly in the hippocampus (Hill & Spencer-Segal, 2021), while DHEA is implicated in neurogenesis and neural preservation (Karishma & Herbert, 2002; Sakr et al., 2014). Higher DHEA levels are associated with enhanced performance in executive functions, concentration, and working memory (Davis et al., 2008). Correspondingly, the MT group demonstrated improved cognitive performance, particularly in attention, spatial-cognitive abilities, and executive functions (abstraction, reasoning, and cognitive flexibility).

These results are further nuanced by the improved RSA findings which are linked to greater vagal control of the heart, possibly related with increased parasympathetic activity pre/post MT sessions. Specifically, the results showed a significant increase in RSA values between the beginning and the end of, respectively, the 4th and 15th sessions in the MT group, whilst the active control group did not show any change.

When aligned with the results of the cognitive-neuropsychological tests, these findings converge with previous studies highlighting that high RSA resting state-level predicted enhanced cognitive ability (Morgan et al., 2007), attention, working memory (Hansen et al., 2003) and executive control (Mathewson et al., 2011). Furthermore, increased RSA is associated to a good behavioural reactivity, suggesting that appropriate vagal regulation is correlated with behaviour and emotion regulation (Doussard-Roosevelt et al., 2003). These results could be explained by the multi-component nature of the improvisational MT activities, which mainly involve listening to music, singing, and playing instruments, combined with the affect-matched hypothesis. For instance, various studies have identified a positive relationship between music listening (Nadar et al., 2014; Mir et al., 2021; Trappe and Voit, 2016), singing (Somayaji et al., 2022; Tanzmeister et al., 2022), and affect-matching music (De la Torre-Luque et al., 2017), demonstrating attenuated systolic blood pressure, improved peripheral vascular function, increased HRV levels, and enhanced parasympathetic activities.

Regarding the RSA results, it appears that the MT employed in this study, has a parasympathetic influence that could reflect a relaxation phenomenon in participants' bodies (Patriquin et al., 2013). The cortisol/DHEA levels are affected, which in the long term may influence participants' cognition, mood and quality of life (Fancourt & Williamon, 2016; de Menezes et al., 2016; de La Rubia Ortí et al., 2018; Hill & Spencer-Segal, 2021). How could this

be explained? The relationship between the activation of the parasympathetic nervous system (PNS) and increased RSA levels may lead to decreased sympathetic nervous system activity and, consequently, a reduction in cortisol release (Doussard-Roosevelt et al., 2003). Evidence suggests that the PNS regulates the hypothalamic-pituitary-adrenal (HPA) axis by inhibiting the release of corticotropin-releasing hormone and adrenocorticotropic hormone, key regulators of the HPA axis (Thayer & Lane, 2009; Ulrich-Lai & Herman, 2009; Yaribeygi et al., 2017). Increased PNS activity may lead to [i] reduced cortisol levels and [ii] elevated DHEA levels, as seen in previous studies (Bittman, 2001; Umbrello et al., 2019). Therefore, our results suggest that each MT activity has positive "short-term" effects, stimulating the PNS to make participants more relaxed and boost their mood. This, in the long term, may influence the HPA axis, regulating cortisol and DHEA levels, positively affecting participants' cognitive and behavioral domains. It is crucial to note that research on this topic is limited, and further studies are needed to fully understand the relationship between MT and cortisol/DHEA production. However, our findings propose that incorporating MT into care homes' daily routines may offer potential health benefits, including improvements in cortisol/DHEA levels.

The interpretation of our results is linked with the specific improvisational approach used in our MT protocol. This is supported by studies reporting how improvisational music interventions can enhance: different cognitive and social skills (Biasutti & Frezza, 2009; Diaz et al., 2020; Pressing, 1988), psychophysical wellbeing (Bruscia, 1987; MacDonald et al., 2014), educational and developmental aspects of music skills (Biasutti, 2015; Biasutti & Frezza, 2009; Campbell, 2009), self-expression (Koutsoupidou, & Hargreaves, 2009), cognitive and emotional developmental skills (Bruscia, 1988; Bunt & Pavlicevic, 2001; Erkkila et al, 2008; Koutsoupidou, & Hargreaves, 2009), and can alter the secretion of steroid hormones (Fukui & Toyoshima, 2008; Takahasci, & Matsuchita, 2006).

Limitations and Future Developments

Strengths of the present research project are the robustness of the RCT design with a matched active control group, protocol fidelity procedures, and the use of specific quantitative measures in three domains: cognitive-neuropsychological, well-being and biomarkers.

Some limitations should be considered: (1) Caregivers in the project were aware of the intervention type, potentially impacting responses in certain behavioural tests (e.g., NPI, BADL). A cluster RCT, randomizing across care homes instead of within, may address challenges in implementing blinding procedures. This approach could

mitigate contamination effects, such as staff interacting differently with participants based on their specific intervention affiliation; (2) Although we aimed to keep standard care activities separate from the intervention, there is still a concern about subtle contamination effects. Staff interactions with participants could have been unintentionally influenced by the therapeutic aspects of the interventions. Future studies should explore methods to better control and minimize these effects, such as using measures to assess and control for unintended influences on study outcomes. (3) Attrition, common in this population, occurred as some participants withdrew during the intervention. Future studies should consider recruiting a larger number of participants to address such challenges; (4) The use of the MMSE in this study, while practical, may not fully capture mild stages of cognitive impairment, especially in college-educated participants. Future studies should explore the utility of alternative cognitive screening assessments with established and validated grading criteria to more accurately assess varying levels of cognitive impairment across diverse educational backgrounds; (5) Possible medication impact on cognitive-psychophysiological and hormonal levels. Due to the sample nature (older adults with cognitive decline in care homes), all participants were on medications, with 76% using cardiovascular medications. A comprehensive analysis excluding those on medications with chronotropic effects was not feasible; (6) The use of only one active control group may limit the generalisability of the results; (7) Due to the nature of the interventions, separate fidelity checklists were used for music therapists and storytellers (see section *SM3* and *SM6* of the *supplementary material)*, which may have hindered direct comparison of fidelity between the two groups. While efforts were made to ensure adherence to protocol within each intervention group, the use of different checklists could potentially introduce variability in fidelity assessment. Future studies could consider standardized fidelity assessment tools across different intervention types to facilitate clearer comparison and ensure consistency in fidelity monitoring. (8) Additionally, it should be noted that the charity funding this research has a well-established relationship with several universities, primarily supporting research in arts therapies, including the MT field. This relationship may influence the selection and training processes of both music therapists and activity coordinators who administered the interventions. While this ensures that staff members are highly trained and proficient, it could also introduce biases or perceptions of bias due to their close affiliation with the charity or their varying levels of training. Future studies should consider involving external therapists and activity coordinators to provide a broader perspective and mitigate potential biases; (9) Furthermore, as no follow-up has been conducted, our understanding of the potential prolonged effects of the interventions is limited.

This study sets the groundwork for future research to explore differences among one-to-one MT, small-group MT, and community MT for participants with varying cognitive abilities. It suggests further investigations combining cognitive-neuropsychological and behavioural tests with a broader array of biomarkers (e.g., apolipoprotein, oxytocin, microRNA) to understand the mechanisms underlying MT-induced changes. Subsequent studies shold delve into the effects of diverse music-based interventions (e.g., improvisational vs. other MT approaches, forms of music training like BAPNE, Romero-Naranjo, 2014) by examining both neuropsychological and biological measures. Additionally, research exploring the potential societal impact of incorporating MT interventions into healthy aging preventative programs to mitigate cognitive decline is warranted.

Conclusion

The study demonstrated that sixteen sessions of one-to-one interactive improvisational MT, delivered over five months, positively impacted older adults with mild to moderate cognitive decline in care homes, compared to one-to-one ST activities. This suggests the potential utility of MT as a non-invasive, non-pharmacological therapy in medical care, promoting brain plasticity and enhancing cognitive-physiological functions, mood, and wellbeing. The positive biomarker outcomes (RSA and cortisol/DHEA) suggest exploring MT as a complementary approach in the potential prevention of neurocognitive diseases (Naftolin et al., 2018).

References

Aalten, P., Verhey, F. R. J., Boziki, M., Bullock, R., Byrne, E. J., Camus, V., Caputo, M., Collins, D., Deyn, P. P. de, Elina, K., Frisoni, G., Girtler, N., Holmes, C., Hurt, C., Marriott, A., Mecocci, P., Nobili, F., Ousset, P. J., Reynish, E., … Robert, P. H. (2007). Neuropsychiatric syndromes in dementia. Results from the European alzheimer disease consortium: part I. *Dementia and Geriatric Cognitive Disorders*, *24*(6), 457–463. https://doi.org/10.1159/000110738

Aki, Ö. E., Alkan, B., Demirsöz, T., Velİbaşoğlu, B., Taşdemir, T., Erbaş, S. P., ... & Demİr, B. (2022). Effects of age, gender and education on phonemic and semantic verbal fluency. *Turk psikiyatri dergisi*, *33*(1), 53. https://doi.org/10.5080/u25553

Aktas, A., & Keskin, B. (2013). Statistical power analysis. *The 7th International Days of Statistics and Economics*, 578–587. http://msed.vse.cz/files/2013/224-Keskin-Burak-paper.pdf

Alexopoulos, G. S., Abrams, R. C., Young, R. C., & Shamoian, C. A. (1988). Cornell scale for depression in dementia. *Biological Psychiatry*, *23*(3), 271–284. https://doi.org/10.1016/0006-3223(88)90038-8

Almeida, D. M., Piazza, J., Liu, Y., & Zarit, S. H. (2020). Use of saliva to better understand the daily experience of adulthood and aging. *Salivary Bioscience: Foundations of Interdisciplinary Saliva Research and Applications*, 655-674. https://doi.org/10.1007/978-3-030-35784-9_27

Altenmüller, E., & Schlaug, G. (2015). Apollo's gift: new aspects of neurologic music therapy. *Progress in Brain Research*, *217*, 237–252. https://doi.org/10.1016/BS.PBR.2014.11.029

An, R., & Liu, G. G. (2016). Cognitive impairment and mortality among the oldest-old Chinese. *International journal of geriatric psychiatry*, *31*(12), 1345-1353. https://doi.org/10.1002/gps.4442

Aoki, M., Okuda, H., Ishihara, H., Hayashi, H., Ohashi, T., Nishihori, T., & Kuze, B. (2020). Hearing loss is associated with hippocampal atrophy and high cortisol/dehydroepiandrosterone sulphate ratio in older adults. *International Journal of Audiology*, *60*(4), 293-299. https://Doi.Org/10.1080/14992027.2020.1831703

Ardila, A., Ostrosky-Solís, F., & Bernal, B. (2006). Cognitive testing toward the future: The example of Semantic Verbal Fluency (ANIMALS). *International Journal of Psychology*, *41*(5), 324–332. https://doi.org/10.1080/00207590500345542

Batubara, S. O., Saragih, I. D., Mulyadi, M., & Lee, B. O. (2023). Effects of art therapy for people with mild or major neurocognitive disorders: a systematic review and meta-analysis. *Archives of Psychiatric Nursing*. https://doi.org/10.1016/j.apnu.2023.04.011

Bassett, J. R. , Marshall, P. M. , & Spillane, R. (1987). The physiological measurement of acute stress (public speaking) in bank employees. *International Journal of Psychophysiology*, *5*(4), 265–273.

Beaty, R. E. (2015). The neuroscience of musical improvisation. *Neuroscience & Biobehavioral Reviews*, *51*, 108–117. https://doi.org/10.1016/j.neubiorev.2015.01.004

Beauchaine, T. (2001). Vagal tone, development, and gray's motivational theory: toward an integrated model of autonomic nervous system functioning in psychopathology. *Development and Psychopathology*, *13*(2), 183–214. https://doi.org/10.1017/S0954579401002012

Berston, G. G., Thomas, B. J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., Nagaraja, H. N., Porges, S. W., Saul, J. P., Stone, P. H., & van der Molen, M. W. (1997). Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology*, *34*(6), 623–648. https://doi.org/10.1111/j.1469-8986.1997.tb02140.x

Bhattarai, K. R., Kim, H. R., & Chae, H. J. (2018). Compliance with saliva collection protocol in healthy volunteers: strategies for managing risk and errors. *International Journal of Medical Sciences*, *15*(8), 823. https://doi.org/10.7150/IJMS.25146

Bialystok, E., & DePape, A.-M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 565–574. https://doi.org/10.1037/a0012735

Biasutti, M. (2017). Teaching improvisation through processes. Applications in music education and implications for general education. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.00911

Biasutti, M., & Frezza, L. (2009). Dimensions of music improvisation. *Creativity Research Journal*, *21*(2–3), 232–242. https://doi.org/10.1080/10400410902861240

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

Bittman, B. B., Berk, L. S., Felten, D. L., & Westengard, J. (2001). Composite effects of group drumming music therapy on modulation of neuroendocrine-immune parameters in normal subjects. *Alternative Therapies in Heath and Medicine*, *7*(1), 38. https://search.proquest.com/openview/0a96c90d6073ac3ed8e97f74f71f964c/1?pq-origsite=gscholar&cbl=32528

Boyd, P. A., Wilks, S. E., & Geiger, J. R. (2018). Activities of Daily Living Assessment among Nursing Home Residents with Advanced Dementia: Psychometric Reevaluation of the Bristol Activities of Daily Living Scale. Health & Social Work, 43(2), 101–108. https://doi.org/doi:10.1093/hsw/hly010

Bruscia, K. E. (1987). Improvisation assessment profiles - The Bruscia model. *Improvisational Models of Music Therapy*, 401–487.

Buasi, N., & Permsuwan, U. (2014). Validation of the Thai QOL-AD version in Alzheimer's patients and caregivers. *The Australasian medical journal*, *7*(6), 251. https://doi.org/10.4066/AMJ.2014.2078

Bucks, R. S., Ashworth, D. L., Wilcock, G. K., & Siegfried, K. (1996). Assessment of activities of daily living in dementia: development of the Bristol activities of daily living scale. *Age and Ageing*, *25*(2), 113–120. https://doi.org/10.1093/ageing/25.2.113

Bugos, J. A., Perlstein, W. M., McCrae, C. S., Brophy, T. S., & Bedenbaugh, P. H. (2007). Individualized piano instruction enhances executive functioning and working memory in older adults. *Aging & Mental Health*, *11*(4), 464–471. https://doi.org/10.1080/13607860601086504

Bush, L. K., Hess, U., & Wolford, G. (1993). Transformations for within-subject designs: A monte carlo investigation. *Psychological Bulletin*, *113*(3), 566–579. https://doi.org/10.1037/0033-2909.113.3.566

Butler, E. A., Wilhelm, F. H., & Gross, J. J. (2006). Respiratory sinus arrhythmia, emotion, and emotion regulation during social interaction. *Psychophysiology*, *43*(6), 612–622. https://doi.org/10.1111/j.1469-8986.2006.00467.x

Chen, C. C. G., & Parker, C. R. (2004). Adrenal androgens and the immune system. *Seminars in Reproductive Medicine*, *22*(04), 369–377. https://doi.org/10.1055/s-2004-861553

Chu, H., Yang, C. Y., Lin, Y., Ou, K. L., Lee, T. Y., O'Brien, A. P., & Chou, K. R. (2014). The impact of group music therapy on depression and cognition in elderly persons with dementia: a randomized controlled study. *Biological Research for Nursing*, *16*(2), 209–217. https://doi.org/10.1177/1099800413485410

Chu, H., Yang, C.-Y., Lin, Y., Ou, K.-L., Lee, T.-Y., O'Brien, A. P., & Chou, K.-R. (2013). The impact of group music therapy on depression and cognition in elderly persons with dementia: a randomized controlled study. *Http://Dx.Doi.Org/10.1177/1099800413485410*, *16*(2), 209–217. https://doi.org/10.1177/1099800413485410

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.

Critchley, M. (1953). *The Parietal Lobes*. Hafner, New York.

Cummings, J. L., Mega, M., Gray, K., Rosenberg-Thompson, S., Carusi, D. A., & Gornbein, J. (1994). The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology*, *44*(12), 2308–2308. https://doi.org/10.1212/WNL.44.12.2308

Darrow, A. A. (2008). *Introduction to approaches in music therapy.* (American Music Therapy Association, Ed.). ERIC Home.

Davis, S. R., Shah, S. M., McKenzie, D. P., Kulkarni, J., Davison, S. L., & Bell, R. J. (2008). Dehydroepiandrosterone sulfate levels are associated with more favorable cognitive function in women. *The Journal of Clinical Endocrinology & Metabolism*, *93*(3), 801–808. https://doi.org/10.1210/jc.2007-2128

de La, J. E., Ortí, R., Pilar, M., Pardo, G., Pérez-Ros, P., Rochina, M. J., & Castillo, S. S. (2018). Iga and alpha-amylase: new targets for well-being in alzheimers disease patients, a pilot study with music therapy. *Neuropsychiatry*, *8*(1), 378–383. https://doi.org/10.4172/NEUROPSYCHIATRY.1000359

de La Rubia Ortí, J. E., García-Pardo, M. P., Iranzo, C. C., Madrigal, J. J. C., Castillo, S. S., Rochina, M. J., & Gascó, V. J. P. (2018). Does music therapy improve anxiety and depression in alzheimer's patients? *Journal of Alternative and Complementary Medicine (New York, N.Y.)*, *24*(1), 33–36. https://doi.org/10.1089/ACM.2016.0346

de Menezes, K. J., Peixoto, C., Nardi, A. E., Carta, M. G., Machado, S., & Veras, A. B. (2016). Dehydroepiandrosterone, its sulfate and cognitive functions. *Clinical Practice & Epidemiology in Mental Health*, *12*(1), 24–37. https://doi.org/10.2174/1745017901612010024

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

Diaz Abrahan, V. M., Shifres, F., & Justel, N. R. (2020). Music improvisation modulates emotional memory. *Psychology of Music*, *48*(4), 465–479. https://doi.org/10.1177/0305735618810793

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Drinkwater, C., Wildman, J., & Moffatt, S. (2019). Social prescribing. *Bmj*, *364*(l1285). https://doi.org/10.1136/bmj.l1285

Doussard-Roosevelt, J. A., Montgomery, L. A., & Porges, S. W. (2003). Short-term stability of physiological measures in kindergarten children: Respiratory sinus arrhythmia, heart period, and cortisol. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *43*(3), 230-242. https://doi.org/10.1002/dev.10136

Dyer, S. M., Harrison, S. L., Laver, K., Whitehead, C., & Crotty, M. (2018). An overview of systematic reviews of pharmacological and non-pharmacological interventions for the treatment of behavioral and psychological symptoms of dementia. *International Psychogeriatrics*, *30*(3), 295–309. https://doi.org/10.1017/S1041610217002344

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers 1996 28:1*, *28*(1), 1–11. https://doi.org/10.3758/BF03203630

Fancourt, D., Williamon, A., Carvalho, L. A., Steptoe, A., Dow, R., & Lewis, I. (2016). Singing modulates mood, stress, cortisol, cytokine and neuropeptide activity in cancer patients and carers. *Ecancermedicalscience*, *10*. https://doi.org/10.3332/ecancer.2016.631

Freitas, S., Simões, M. R., Marôco, J., Alves, L., & Santana, I. (2012). Construct validity of the montreal cognitive assessment (MoCA). *Journal of the International Neuropsychological Society*, *18*(2), 242-250. https://doi.org/10.1017/S1355617711001573

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, 12(3), 189-198.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/10.1177/0305735614548500

Franzin, A., Sambo, F., & Di Camillo, B. (2017). bnstruct: an R package for Bayesian Network structure learning in the presence of missing data. *Bioinformatics*, 33(8), 1250-1252. https://doi.org/10.1093/bioinformatics/btw807

Frazier, T. W., Strauss, M. E., & Steinhauer, S. R. (2004). Respiratory sinus arrhythmia as an index of emotional response in young adults. *Psychophysiology*, *41*(1), 75–83. https://doi.org/10.1046/j.1469-8986.2003.00131.x

Friedman, N., Nachman, I., & Pe'er, D. (2013). Learning Bayesian network structure from massive datasets: The "sparse candidate" algorithm. *arXiv preprint* arXiv:1301.6696. https://doi.org/10.48550/arXiv.1301.6696

Frye, C. A., & Lacey, E. H. (1999). The neurosteroids DHEA and DHEAS may influence cognitive performance by altering affective state. *Physiology & Behavior*, *66*(1), 85–92. https://doi.org/10.1016/S0031-9384(98)00256-X

Fukui, H., & Toyoshima, K. (2008). Music facilitate the neurogenesis, regeneration and repair of neurons. *Medical Hypotheses*, *71*(5), 765–769. https://doi.org/10.1016/j.mehy.2008.06.019

Fusar-Poli, L., Bieleninik, Ł., … N. B.-A. & mental, & 2018, undefined. (2017). The effect of music therapy on cognitive functions in patients with dementia: a systematic review and meta-analysis. *Taylor & Francis*, *22*(9), 1097–1106. https://doi.org/10.1080/13607863.2017.1348474

Gauthier, S., Rosa-Neto, P., Morais, J. A., & Webster, C. (2021). World Alzheimer Report 2021: journey through the diagnosis of dementia*. Alzheimer's Disease International*. https://www.alzint.org/resource/world-alzheimer-report-2021/

Gilsanz, P., Quesenberry, C. P., Mayeda, E. R., Glymour, M. M., Farias, S. T., & Whitmer, R. A. (2019). Stressors in midlife and risk of dementia. *Alzheimer Disease & Associated Disorders*, *33*(3), 200–205. https://doi.org/10.1097/WAD.0000000000000313

Grover, S., Sandhu, P., Nijjar, G. S., Percival, A., Chudyk, A. M., Liang, J., ... & Ashe, M. C. (2023). Older adults and social prescribing experience, outcomes, and processes: a meta-aggregation systematic review. *Public health,* 218, 197-207. https://doi.org/10.1016/j.puhe.2023.02.016

Grossman, P., Beek, J., & Wientjes, C. (1990). A comparison of three quantification methods for estimation of respiratory sinus arrhythmia. *Psychophysiology*, *27*(6), 702–714. https://doi.org/10.1111/j.1469-8986.1990.tb03198.x

Grossman, P., & Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology*, *74*(2), 263–285. https://doi.org/10.1016/J.BIOPSYCHO.2005.11.014

Hansen, A. L., Johnsen, B. H., & Thayer, J. F. (2003). Vagal influence on working memory and attention. *International Journal of Psychophysiology*, *48*(3), 263–274. https://doi.org/10.1016/S0167-8760(03)00073-4

Harrell, L. E., Marson, D., Chatterjee, A., & Parrish, J. A. (2000). The severe mini-mental state examination: a new neuropsychologic instrument for the bedside assessment of severely impaired patients with Alzheimer disease. *Alzheimer Disease & Associated Disorders*, *14*(3), 168-175.

Hegde, S. (2014). Music-based cognitive remediation therapy for patients with traumatic brain injury. *Frontiers in Neurology*, *5*. https://doi.org/10.3389/fneur.2014.00034

Helmes, E., Csapo, K. G., & Short, J. -a. (1987). Standardization and validation of the multidimensional observation scale for elderly subjects (MOSES). *Journal of Gerontology*, *42*(4), 395–405. https://doi.org/10.1093/geronj/42.4.395

Hildreth, K. L., Gozansky, W. S., Jankowski, C. M., Grigsby, J., Wolfe, P., & Kohrt, W. M. (2013). Association of serum dehydroepiandrosterone sulfate and cognition in older adults: Sex steroid, inflammatory, and metabolic mechanisms. *Neuropsychology*, *27*(3), 356–363. https://doi.org/10.1037/a0032230

Hill, A. R., & Spencer-Segal, J. L. (2021). Glucocorticoids and the brain after critical illness. *Endocrinology*, *162*(3). https://doi.org/10.1210/endocr/bqaa242

Huang, Y. Y., Teng, T., Giovane, C. D., Wang, R. Z., Suckling, J., Shen, X. N., ... & Yu, J. T. (2023). Pharmacological treatment of neuropsychiatric symptoms of dementia: a network meta-analysis. *Age and ageing*, 52(6), 1-15. https://doi.org/10.1093/ageing/afad091

Hudak, E. M., Bugos, J., Andel, R., Lister, J. J., Ji, M., & Edwards, J. D. (2019). Keys to staying sharp: a randomized clinical trial of piano training among older adults with and without mild cognitive impairment. *Contemporary clinical trials*, *84*, 105789. https://doi.org/10.1016/j.cct.2019.06.003

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 randomised controlled feasibility study. *BMC Geriatrics*, *15*(1), 1–19. https://doi.org/10.1186/s12877-015-0082-4

Jensen, F. V. (1996). *An introduction to Bayesian networks* (Vol. 210, pp. 1-178). London: UCL press.

Ji, E., Weickert, C. S., Purves-Tyson, T., White, C., Handelsman, D. J., Desai, R., O'Donnell, M., Liu, D., Galletly, C., Lenroot, R., & Weickert, T. W. (2021). Cortisol-dehydroepiandrosterone ratios are inversely associated with hippocampal and prefrontal brain volume in schizophrenia. *Psychoneuroendocrinology*, *123*, 104916. https://doi.org/10.1016/J.PSYNEUEN.2020.104916

Jin, R. O., Mason, S., Mellon, S. H., Epel, E. S., Reus, V. I., Mahan, L., Rosser, R. L., Hough, C. M., Burke, H. M., Mueller, S. G., & Wolkowitz, O. M. (2016). Cortisol/DHEA ratio and hippocampal volume: a pilot study in major depression and healthy controls. *Psychoneuroendocrinology*, *72*, 139–146. https://doi.org/10.1016/J.PSYNEUEN.2016.06.017

Jünemann, K., Marie, D., Worschech, F., Scholz, D. S., Grouiller, F., Kliegel, M., ... & Sinke, C. (2022). Six months of piano training in healthy elderly stabilizes white matter microstructure in the fornix, compared to an active control group. *Frontiers in aging neuroscience*, 14, 817889.

Kalmijn, S., Launer, L. J., Stolk, R. P., de Jong, F. H., Pols, H. A. P., Hofman, A., Breteler, M. M. B., & Lamberts, S. W. J. (1998). A prospective study on cortisol, dehydroepiandrosterone sulfate, and cognitive function in the elderly. *The Journal of Clinical Endocrinology & Metabolism*, *83*(10), 3487–3492. https://doi.org/10.1210/JCEM.83.10.5164

Karishma, K. K., & Herbert, J. (2002). Dehydroepiandrosterone (DHEA) stimulates neurogenesis in the hippocampus of the rat, promotes survival of newly formed neurons and prevents corticosterone-induced suppression. *European Journal of Neuroscience*, *16*(3), 445–453. https://doi.org/10.1046/j.1460-9568.2002.02099.x

Keevil, B. G. (2013). Novel liquid chromatography tandem mass spectrometry (LC-MS/MS) methods for measuring steroids. *Best practice & research Clinical endocrinology & metabolism*, *27*(5), 663-674.

Kim, S. J., & Yoo, G. E. (2019). Instrument playing as a cognitive intervention task for older adults: a systematic review and meta-analysis. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00151

Kimonides, V. G. , Spillantini, M. G. , Sofroniew, M. V. , Fawcett, J. W. , & & Herbert, J. (1999). Dehydroepiandrosterone antagonizes the neurotoxic effects of corticosterone and translocation of stress-activated protein kinase 3 in hippocampal primary cultures. *Neuroscience*, *89*(2), 429–436.

Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: principles and techniques*. MIT press.

Kurata, K., Takebayashi, M., Morinobu, S., & Yamawaki, S. (2004). β-Estradiol, dehydroepiandrosterone, and dehydroepiandrosterone sulfate protect against *n* -methyl-d-aspartate-induced neurotoxicity in rat hippocampal neurons by different mechanisms. *Journal of Pharmacology and Experimental Therapeutics*, *311*(1), 237–245. https://doi.org/10.1124/jpet.104.067629

Laver, K., Dyer, S., Whitehead, C., Clemson, L., & Crotty, M. (2016). Interventions to delay functional decline in people with dementia: a systematic review of systematic reviews. *BMJ Open* , *6*(4).

Leggieri, M., Thaut, M. H., Fornazzari, L., Schweizer, T. A., Barfett, J., Munoz, D. G., & Fischer, C. E. (2019). Music Intervention Approaches for alzheimer's disease: a review of the literature. *Frontiers in Neuroscience*, *0*, 132. https://doi.org/10.3389/FNINS.2019.00132

Lennartsson, A. K., Theorell, T., Kushnir, M. M., Bergquist, J., & Jonsdottir, I. H. (2013). Perceived stress at work is associated with attenuated DHEA-S response during acute psychosocial stress. *Psychoneuroendocrinology*, *38*(9), 1650–1657. https://doi.org/10.1016/J.PSYNEUEN.2013.01.010

Lennartsson, A. K., Theorell, T., Rockwood, A. L., Kushnir, M. M., & Jonsdottir, I. H. (2013). Perceived stress at work is associated with lower levels of DHEA-S. *PloS One*, *8*(8). https://doi.org/10.1371/JOURNAL.PONE.0072460

Li, H., Klein, G. , Sun, P. , & & Buchan, A. M. (2001). Dehydroepiandrosterone (DHEA) reduces neuronal injury in a rat model of global cerebral ischemia. *Brain Research*, *888*(2), 263–266.

Li, H. C., Wang, H. H., Lu, C. Y., Chen, T. B., Lin, Y. H., & Lee, I. (2019). The effect of music therapy on reducing depression in people with dementia: A systematic review and meta-analysis. *Geriatric Nursing*, *40*(5), 510–516. https://doi.org/10.1016/J.GERINURSE.2019.03.017

Lin, T. H., Liao, Y. C., Tam, K. W., Chan, L., & Hsu, T. H. (2023). Effects of music therapy on cognition, quality of life, and neuropsychiatric symptoms of patients with dementia: a systematic review and meta-analysis of randomized controlled trials. *Psychiatry Research*, 115498. https://doi.org/10.1016/j.psychres.2023.115498

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

Logsdon, R. G., Gibbons, L. E., McCurry, S. M., & Teri, L. (2002). Assessing quality of life in older adults with cognitive impairment. *Psychosomatic Medicine*, *64*(3), 510–519. https://journals.lww.com/psychosomaticmedicine/toc/2002/05000

López-Ortega, M., Torres-Castro, S., & Rosas-Carrasco, O. (2016). Psychometric properties of the Satisfaction with Life Scale (SWLS): secondary analysis of the Mexican Health and Aging Study. Health and Quality of Life Outcomes, 14(1). . https://doi.org/doi:10.1186/s12955-016-0573-9

Lü, W., & Wang, Z. (2018). Associations between resting respiratory sinus arrhythmia, intraindividual reaction time variability, and trait positive affect. *Emotion*, *18*(6), 834.

Lupien, S. J., de Leon, M., de Santi, S., Convit, A., Tarshish, C., Nair, N. P. v., Thakur, M., McEwen, B. S., Hauger, R. L., & Meaney, M. J. (1998). Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nature Neuroscience*, *1*(1), 69–73. https://doi.org/10.1038/271

Mangiacotti, 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

Mathewson, K. J., Drmic, I. E., Jetha, M. K., Bryson, S. E., Goldberg, J. O., Hall, G. B., Santesso, D. L., Segalowitz, S. J., & Schmidt, L. A. (2011). Behavioral and cardiac responses to emotional stroop in adults with autism spectrum disorders: influence of medication. *Autism Research*, *4*(2), 98–108. https://doi.org/10.1002/aur.176

McDermott, O., Crellin, N., … H. R.-I. journal of, & 2013, undefined. (2012). Music therapy in dementia: a narrative synthesis systematic review. *Wiley Online Library*, *28*(8), 781–794. https://doi.org/10.1002/gps.3895

McPherson, G. E. (2005). From child to musician: skill development during the beginning stages of learning an instrument. *Psychology of Music*, *33*(1), 5–35. https://doi.org/10.1177/0305735605048012

Michael, A., Jenaway, A., Paykel, E. S., & Herbert, J. (2000). Altered salivary dehydroepiandrosterone levels in major depression in adults. *Biological Psychiatry*, *48*(10), 989–995. https://doi.org/10.1016/S0006-3223(00)00955-0

Mitchell, A.J. (2013*). The mini-mental state examination (MMSE): an update on its diagnostic validity for cognitive disorders*. In: Larner, A. (eds) Cognitive Screening Instruments. Springer, London. https://doi.org/10.1007/978-1-4471-2452-8_2

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.

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, *22*(11), 1425–1433. https://doi.org/10.1177/0956797611416999

Moreno-Morales, C., Calero, R., Moreno-Morales, P., & Pintado, C. (2020). Music therapy in the treatment of dementia: a systematic review and meta-analysis. *Frontiers in Medicine*, *0*, 160. https://doi.org/10.3389/FMED.2020.00160

Morgan, C. A., Aikins, D. E., Steffian, G., Coric, V., & Southwick, S. (2007). Relation between cardiac vagal tone and performance in male military personnel exposed to high stress: three prospective studies. *Psychophysiology*, *44*(1). https://doi.org/10.1111/j.1469-8986.2006.00475.x

Morgan, C. A., Southwick, S., Hazlett, G., Rasmusson, A., Hoyt, G., Zimolo, Z., & Charney, D. (2004). Relationships among plasma dehydroepiandrosterone sulfate and cortisol levels, symptoms of dissociation, and objective performance in humans exposed to acute stress. *Archives of General Psychiatry*, *61*(8), 819–825. https://doi.org/10.1001/ARCHPSYC.61.8.819

Naftolin, F., Silva, I., & Orley, A. (2018). *Reproductive Hormones and Dementia* (pp. 191–201). https://doi.org/10.1007/978-3-319-78009-2_16

Nasreddine, Z. S., Phillips, N. A., BÃ©dirian, 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

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.

Patriquin, M. A., Scarpa, A., Friedman, B. H., & Porges, S. W. (2013). Respiratory sinus arrhythmia: a marker for positive social functioning and receptive language skills in children with autism spectrum disorders. *Developmental Psychobiology*, *55*(2), 101–112. https://doi.org/10.1002/dev.21002

Pavot, W., & Diener, E. (2009). *Review of the Satisfaction With Life Scale* (pp. 101–117). https://doi.org/10.1007/978-90-481-2354-4_5

Polenick, C. A., Birditt, K. S., Turkelson, A., Perbix, E. A., Salwi, S. M., & Zarit, S. H. (2021). Daily social interactions and HPA axis activity among midlife and older adults. *The Gerontologist*, *61*(6), 897-906. https://doi.org/10.1093/geront/gnaa215

Puyvelde, M. van, Loots, G., Vanfleteren, P., Meys, J., Simcock, D., & Pattyn, N. (2014). Do you hear the same? Cardiorespiratory responses between mothers and infants during tonal and atonal music. *PLOS ONE*, *9*(9), e106920. https://doi.org/10.1371/JOURNAL.PONE.0106920

Reitan, R. M., & Wolfson, D. (1995). Category test and trail making test as measures of frontal lobe functions. *The Clinical Neuropsychologist*, *9*(1), 50–56. https://doi.org/10.1080/13854049508402057

Reuther, S., Dichter, M. N., Bartholomeyczik, S., Nordheim, J., & Halek, M. (2016). Construct validity and internal consistency of the neuropsychiatric inventory – nursing home (NPI-NH) in German nursing homes. International Psychogeriatrics, 28(06), 1017–1027. https://doi.org/doi:10.1017/s1041610215002343

Rey A (1958). *L'examen clinique en psychologie (1st ed.)*. Paris: Presses universitaires de France.

Sakr, H. F. , Khalil, K. I. , Hussein, A. M. , Zaki, M. S. , Eid, R. A. , & & Alkhateeb, M. (2014). Effect of dehydroepiandrosterone (DHEA) on memory and brain derived neurotrophic factor (BDNF) in a rat model of vascular dementia. *J Physiol Pharmacol*, *65*(1), 41–53.

Savva, G. M., Zaccai, J., Matthews, F. E., Davidson, J. E., McKeith, I., & Brayne, C. (2009). Prevalence, correlates and course of behavioural and psychological symptoms of dementia in the population. *The British Journal of Psychiatry*, 194(3), 212-219. https://doi.org/10.1192/bjp.bp.108.049619

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341–349. https://doi.org/10.1111/1469-8986.00172.x

Shi, S., Mörike, K., & Klotz, U. (2008). The clinical implications of ageing for rational drug therapy. *European Journal of Clinical Pharmacology*, *64*(2), 183–199. https://doi.org/10.1007/S00228-007-0422-1

Sittler, M. C., Worschech, F., Wilz, G., Fellgiebel, A., & Wuttke-Linnemann, A. (2021). Psychobiological mechanisms underlying the health-beneficial effects of music in people living with dementia: A systematic review of the literature. *Physiology & Behavior*, *233*, 113338. https://doi.org/10.1016/J.PHYSBEH.2021.113338

Skoldunger, A., Johnell, K., Winblad, B., & Wimo, A. (2013). Mortality and treatment costs have a great impact on the cost- effectiveness of disease modifying treatment in alzheimer's disease - a simulation study. *Current Alzheimer Research*, *10*(2), 207–2013.

Spreen, O. & Strauss, E (1998). *A Compendium of Neuropsychological Tests. Administration, Norms, and Commentary.* Oxford University Press, New York.

Stahl, F., & Dörner, G. (1982). Responses of salivary cortisol levels to stress-situations. *Endokrinologie*, *80*(2), 158–162.

Staton, L., El-Sheikh, M., & Buckhalt, J. A. (2009). Respiratory sinus arrhythmia and cognitive functioning in children. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *51*(3), 249-258.

Strong, J. V., & Midden, A. (2020). Cognitive differences between older adult instrumental musicians: Benefits of continuing to play. *Psychology of Music, 48*(1), 67-83. https://doi.org/10.1177/0305735618785020

Suzuki, M., Kanamori, M., Watanabe, M., Nagasawa, S., Kojima, E., Ooshiro, H., & Nakahara, D. (2004). Behavioral and endocrinological evaluation of music therapy for elderly patients with dementia. *Nursing & Health Sciences*, *6*(1), 11–18. https://doi.org/10.1111/J.1442-2018.2003.00168.X

Swaine, J. S. (2014). Musical communication, emotion regulation and the capacity for attention control: a theoretical model. *Psychology of Music*, *42*(6), 856–863. https://doi.org/10.1177/0305735614545197

Tafet, G. E., Idoyaga-Vargas, V. P., Abulafia, D. P., Calandria, J. M., Roffman, S. S., Chiovetta, A., & Shinitzky, M. (2001). Correlation between cortisol level and serotonin uptake in patients with chronic stress and depression. *Cognitive, Affective, & Behavioral Neuroscience*, *1*(4), 388–393. https://doi.org/10.3758/CABN.1.4.388

Takahashi, T., & Matsushita, H. (2006). Long-term effects of music therapy on elderly with moderate/severe Dementia. *Journal of Music Therapy*, *43*(4), 317–333. https://doi.org/10.1093/JMT/43.4.317

Tamagno, E., Guglielmotto, M., Bardini, P., Santoro, G., Davit, A., di Simone, D., Danni, O., & Tabaton, M. (2003). Dehydroepiandrosterone reduces expression and activity of BACE in NT2 neurons exposed to oxidative stress. *Neurobiology of Disease*, *14*(2), 291–301. https://doi.org/10.1016/S0969-9961(03)00131-1

Tannenbaum, C., Barrett-Connor, E., Laughlin, G. A., & Platt, R. W. (2003). A longitudinal study of dehydroepiandrosterone sulphate (DHEAS) change in older men and women: The Rancho Bernardo Study. *European Journal of Endocrinology*, *151*(6), 717–725. https://doi.org/10.1530/eje.0.1510717

Thayer, J. F., & Lane, R. D. (2000). A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of Affective Disorders*, *61*(3), 201–216. https://doi.org/10.1016/S0165-0327(00)00338-4

Thayer, J. F., & Lane, R. D. (2007). The role of vagal function in the risk for cardiovascular disease and mortality. *Biological Psychology*, *74*(2), 224–242. https://doi.org/10.1016/j.biopsycho.2005.11.013

Thayer, J. F., & Lane, R. D. (2009). Claude Bernard and the heart–brain connection: Further elaboration of a model of neurovisceral integration. *Neuroscience & Biobehavioral Reviews*, *33*(2), 81–88. https://doi.org/10.1016/j.neubiorev.2008.08.004

Theorell, T., Engströ, G., Hallinder, H., Lennartsson, A.-K., Kowalski, J., & Emami, A. (2021). The use of saliva steroids (cortisol and DHEA-s) as biomarkers of changing stress levels in people with dementia and their caregivers: A pilot study. *Article SCIENCE PROGRESS Science Progress*, *104*(2), 1–18. https://doi.org/10.1177/00368504211019856

Tian, C. H. E. N., Manfei, X. U., Justin, T. U., Hongyue, W. A. N. G., & Xiaohui, N. I. U. (2018). Relationship between Omnibus and Post-hoc Tests: An Investigation of performance of the F test in ANOVA. *Shanghai archives of psychiatry*, *30*(1), 60. 10.11919/j.issn.1002-0829.218014

Tombaugh, T. N., & McIntyre, N. J. (1992). The Mini-Mental State Examination: a comprehensive review. *Journal of the American Geriatrics Society*, *40*(9), 922–935. https://doi.org/10.1111/j.1532-5415.1992.tb01992.x

Tsamardinos, I., Brown, L. E., & Aliferis, C. F. (2006). The max-min hill-climbing Bayesian network structure learning algorithm. *Machine learning*, 65, 31-78. https://doi.org/10.1007/s10994-006-6889-7

Ulrich-Lai, Y. M., & Herman, J. P. (2009). Neural regulation of endocrine and autonomic stress responses. *Nature Reviews Neuroscience*, 10(6), 397-409. https://doi.org/10.1038/nrn2647

Umbrello, M., Sorrenti, T., Mistraletti, G., Formenti, P., Chiumello, D., & Terzoni, S. (2019). Music therapy reduces stress and anxiety in critically ill patients: a systematic review of randomized clinical trials. *Minerva anestesiologica*, *85*(8), 886-898.

Van Niekerk, J. K., Huppert, F. A., & Herbert, J. (2001). Salivary cortisol and DHEA: association with measures of cognition and well-being in normal older men, and effects of three months of DHEA supplementation. *Psychoneuroendocrinology*, *26*(6), 591–612. https://doi.org/10.1016/S0306-4530(01)00014-2

Van Puyvelde, M., Loots, G., Vanfleteren, P., Meys, J., Simcock, D., & Pattyn, N. (2014). *Do You Hear the Same? Cardiorespiratory Responses between Mothers and Infants during Tonal and Atonal Music*. https://doi.org/10.1371/journal.pone.0106920

Van Puyvelde, M., Rodrigues, H., Loots, G., De Coster, L., Du Ville, K., Matthijs, L., ... & Pattyn, N. (2014b). Shall we dance? Music as a port of entrance to maternal–infant intersubjectivity in a context of postnatal depression. *Infant mental health journal*, *35*(3), 220-232. https://doi.org/10.1002/imhj.21431

Van Puyvelde, M., Loots, G., Meys, J., Neyt, X., Mairesse, O., Simcock, D., & Pattyn, N. (2015). Whose clock makes yours tick? How maternal cardiorespiratory physiology influences newborns' heart rate variability. *Biological psychology*, *108*, 132-141. https://doi.org/10.1016/j.biopsycho.2015.04.001

Wan, C. Y., & Schlaug, G. (2010). Music making as a tool for promoting brain plasticity across the life span. *The Neuroscientist*, *16*(5), 566–577. https://doi.org/10.1177/1073858410377805

Wang, H.-X., Karp, A., Winblad, B., & Fratiglioni, L. (2002). Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: a longitudinal study from the kungsholmen project. *American Journal of Epidemiology*, *155*(12), 1081–1087. https://doi.org/10.1093/AJE/155.12.1081

Wang, M., Wu, J., & Yan, H. (2023a). Effect of music therapy on older adults with depression: a systematic review and meta-analysis. *Complementary Therapies in Clinical Practice*, 101809. https://doi.org/10.1016/j.ctcp.2023.101809

Wang, Y. Y., Wang, X. X., Chen, L., Liu, Y., & Li, Y. R. (2023b). 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

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

Williams, J. R., & Marsh, L. (2008). Validity of the Cornell scale for depression in dementia in Parkinson's disease with and without cognitive impairment. Movement Disorders, 24(3), 433–437. https://doi:10.1002/mds.22421

World Health Organization. (2017). *Global action plan on the public health response to dementia 2017–2025*. https://apps.who.int/iris/bitstream/handle/10665/259615/?sequence=1

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.

Yao, C. T., Lee, B. O., Hong, H., & Su, Y. C. (2023). Evaluation of the music therapy program interventions on agitated behavior for people with dementia in Taiwan institutional care. *Educational Gerontology*, 49(3), 228-239. https://doi.org/10.1080/03601277.2022.2099076

Yaribeygi, H., Panahi, Y., Sahraei, H., & Johnston, T. P. (2017). The impact of stress on body function: a review. *EXCLI Journal*, 16, 1057-1072. https://doi.org/10.17179/excli2017-480

Yiallouris, A., Tsioutis, C., Agapidaki, E., Zafeiri, M., Agouridis, A. P., Ntourakis, D., & Johnson, E. O. (2019). Adrenal aging and its implications on stress responsiveness in humans. *Frontiers in Endocrinology*, *10*(FEB), 54. https://doi.org/10.3389/FENDO.2019.00054/BIBTEX

Zhang, Y., Cai, J., An, L., Hui, F., Ren, T., Ma, H., & Zhao, Q. (2017). Does music therapy enhance behavioral and cognitive function in elderly dementia patients? A systematic review and meta-analysis. *Ageing Research Reviews*, *35*, 1–11. https://doi.org/10.1016/J.ARR.2016.12.003

Zheng, P. (2009). Neuroactive steroid regulation of neurotransmitter release in the CNS: Action, mechanism and possible significance. *Progress in Neurobiology*, *89*(2), 134–152. https://doi.org/10.1016/j.pneurobio.2009.07.001

Click here to access/download
Supplemental Material for Online Publication
Supplementary material Anonymised2024.pdf