AUREUM ACADEMIC JOURNAL
VOL. 3, NO. 2, June 2026
https://doi.org/10.48161/aaj.v3n2a97

The Impact of Music and Performance-Based Activities on Cognitive Function and Well-Being

Ezeanyi Benson Chukwunonso[1*], Ifeoma Francisca Chukwuma-Nwankwo[2], Eze Ljeoma Ogayi[3]

[1] Science Education, Education, Alex Ekwueme Federal University, Ndufu-Alike, Ebonyi State, Abakaliki, Nigeria;
[2] English and Literary Studies, Faculty of Education, Alex Ekwueme Federal University, Ndufu-Alike, Abakaliki, Nigeria;
[3] Science Education, Federal College of Education, Eha-Amuffu, Alex Ekwueme Federal University, Ndufu-Alike, Ebonyi, Nigeria.

* **Corresponding author:** benson.ezeanyi@funai.edu.ng.

ABSTRACT: Active music-making and performance activities may provide cognitively, emotionally, motorically, and socially demanding experiences in later life, but previous studies rarely separate component effects from attention, expectancy, and group contact. This study estimated the independent effects of music-making and theatre- and movement-based performance, and the uncertainty surrounding their interaction, among adults aged 60–75 years with subjective cognitive decline but no dementia. In a multicentre, assessor-blinded 2×2 factorial randomized trial, 480 participants received music only, performance only, combined delivery, or an attention- and time-matched social-learning program for 16 weeks. Week 16 was the confirmatory endpoint; assessments also occurred at week 8 and at 3, 6, and 12 months. Co-primary outcomes were the NIH Toolbox Fluid Cognition Composite and WEMWBS. Intention-to-treat multiple-membership multilevel models accounted for repeated measures, delivery groups, centers, and exposure-weighted facilitators. Holm correction controlled the four confirmatory tests; multiple imputation, practice-effect sensitivity analyses, and standardized effects assessed robustness. At week 16, music improved fluid cognition by 2.6 points (95% CI 1.1–4.1; Holm-adjusted p=.003) and performance by 2.1 points (95% CI 0.6–3.6; p=.006). Music and performance improved WEMWBS by 2.2 and 2.8 points, respectively. Interaction estimates were imprecise and did not show clear departure from additivity. Effects attenuated after delivery and were uncertain at 12 months. Secondary differences were selective, and longitudinal process analyses remained exploratory and underpowered. Music and performance therefore produced modest short-term cognitive and well-being effects without establishing synergy, durable cognitive protection, or unique causal mechanisms.

Keywords: music-making; performing arts; cognitive aging; subjective cognitive decline; mental well-being.

I. INTRODUCTION

Subjective cognitive decline (SCD) refers to a persistent self-experienced decline in cognition despite performance that remains within broadly normal limits. It is clinically heterogeneous: some adults remain stable, whereas others have an elevated risk of later objective impairment when concern persists and is accompanied by additional risk indicators [1]. This uncertainty creates demand for low-risk interventions that support cognitive and psychological resources without presenting ordinary aging concerns as dementia or implying guaranteed prevention. Contemporary dementia-prevention evidence emphasizes modifiable behavioral and social exposures while rejecting single-activity claims of protection [2].

Participatory arts are plausible because they combine learning, sensory integration, emotion, motivation, movement, and social coordination. Reviews of arts-based randomized studies in older adults report possible cognitive benefits but repeatedly identify small samples, passive comparators, heterogeneous programmes, and short follow-up [3]. Active music participation shows small average cognitive effects in populations with mild impairment or dementia, although psychosocial outcomes vary across studies [4], [5]. Evidence maps similarly show wide variation by dose, delivery form, outcome, and population [6].

Community singing trials illustrate the distinction between promise and certainty. Structured group singing has improved selected subjective and social outcomes [7], while a large community-choir trial reduced loneliness and increased interest in life without uniform improvement across health indicators [8]. A wider review concluded that music and singing can support adult well-being, but effects depend on baseline distress, participant selection, intervention intensity, and comparator strength [9]. These findings justify active rather than passive comparison conditions.

Performance-based activities may recruit partly different capacities. Theatre requires role adoption, narrative sequencing, cue memory, inhibition, perspective shifting, and rapid interpersonal adjustment. Movement-based performance adds spatial updating, balance, motor planning, and embodied timing. Reviews of dance in older adults report possible cognitive benefit but inconsistent executive-function and transfer effects [10], [11]. Senior theatre studies suggest potential psychological and physical value, although controlled cognitive evidence remains sparse [12].

Existing studies rarely determine whether music and performance have independent effects, whether their combination departs from additivity, or whether any change persists after delivery. Parallel multi-arm comparisons are statistically inefficient for component questions, while passive controls leave adult attention, group belonging, routine, expectancy, and facilitator support uncontrolled. The present assessor-blinded 2×2 factorial trial addressed these limitations by delivering music and performance within a constant contact-time architecture and comparing both factors with matched social-learning activity.

The dual co-primary outcomes reflected distinct aims. The NIH Toolbox Fluid Cognition Composite represented standardized cognitive performance, whereas WEMWBS represented positive mental functioning. We kept secondary cognitive, psychological, social, functional, sleep, and quality-of-life outcomes analytically separate. Mechanism measures were temporally ordered but not randomized, and the STAC-r served as an interpretive framework rather than evidence of neural change because the study included no neuroimaging or neurophysiological outcomes.

II. RESEARCH PROBLEM, AIM, OBJECTIVES, AND QUESTIONS

1. RESEARCH PROBLEM

2. RESEARCH AIM AND OBJECTIVES

The study aimed to estimate average factorial effects rather than rank four unrelated arms. Its objectives were: (1) confirmatory estimation of the music main effect and the performance main effect on NIH Toolbox Fluid Cognition and WEMWBS at week 16; (2) secondary estimation of the music × performance interaction and descriptive four-condition means, without assuming synergy or equivalence; (3) evaluation of secondary cognitive, psychological, social, functional, sleep, quality-of-life, implementation, and safety outcomes; and (4) exploratory assessment of temporally ordered process associations and prespecified contextual variation through 12 months.

3. RESEARCH QUESTIONS

- RQ1. What main effects did active group music-making produce in fluid cognition and mental well-being at week 16 and through 12 months?
- RQ2. What main effects did performance-based activity produce in fluid cognition and mental well-being over the same period?
- RQ3. What was the estimated music × performance interaction, how precise was it, and what did the four randomized-condition means show descriptively?

- RQ4. What selective patterns occurred in secondary outcomes, implementation, safety, and temporally ordered exploratory processes?

III. LITERATURE REVIEW AND RESEARCH GAP

1. ACTIVE MUSIC-MAKING: COGNITIVE AND PSYCHOSOCIAL EVIDENCE

Music-making integrates auditory prediction, timed motor response, working memory, error monitoring, sequence learning, and interpersonal coordination. Systematic reviews report small average cognitive effects, with substantial heterogeneity in diagnosis, instrument, dose, and comparator [3]–[5]. Evidence maps show that researchers have studied cognition, mood, quality of life, physiological outcomes, and social functioning, but certainty varies [6]. The strongest community evidence often concerns belonging and positive affect rather than objective cognitive change.

Singing trials show how common intervention elements can influence results. Sing4Health reported changes in selected subjective and social outcomes [7], and Community of Voices reduced loneliness and increased interest in life [8]. Yet both musical content and group membership may contribute. Reviews of adult music and singing programmes conclude that attention, expectancy, facilitator style, baseline distress, and participant preference can alter apparent effects [9]. An attention- and time-matched social-learning comparator is therefore methodologically stronger than wait-list or usual-activity control when the research question concerns added music content.

2. THEATRE, MOVEMENT, AND PERFORMANCE-BASED EVIDENCE

Theatre and movement combine narrative rehearsal, role flexibility, inhibition, memory for cues, embodied sequencing, spatial updating, and cooperative improvisation. Dance reviews in healthy older adults suggest possible cognitive maintenance or improvement, but small trials, inconsistent styles, and heterogeneous endpoints limit precision [10]. Meta-analyses indicate favorable average effects on global cognition and memory, with less consistent findings for executive function [11]. Direct senior-theatre projects report changes in selected depressive and physical outcomes [12], but controlled cognitive evidence remains limited.

In mild cognitive impairment, dance trials show heterogeneity in intensity, motor demand, and risk of bias [13], [14]. Combining theatre and movement into one performance factor can be justified by shared functions—role adaptation, narrative sequencing, embodied communication, and partner response—but it prevents attribution to either component. Consequently, the present study estimated a performance package rather than separate drama and movement effects.

3. SUBJECTIVE COGNITIVE DECLINE, OUTCOME SELECTION, AND PRACTICE EFFECTS

SCD is defined by perceived decline in a person whose objective performance does not meet criteria for mild cognitive impairment or dementia. Subjective concern may reflect anxiety, health beliefs, sleep, sensory loss, normal aging, or preclinical disease. Trials that recruit solely through a screening score risk mixing cognitively normal adults with mild cognitive impairment. Publication-quality studies therefore require explicit criteria, age- and education-sensitive testing, preserved everyday functioning, and independent clinical adjudication.

Outcome selection also influences conclusions. Global screens are useful for eligibility but insensitive to modest domain-specific training effects. Fluid cognition is more responsive to executive, memory, attention, and processing-speed demands, while positive well-being captures a distinct psychosocial aim. Repeated computerized assessment can produce familiarity and strategy effects, particularly in adults with limited digital experience [15]. Therefore, alternate forms, control-group comparisons, practice-effect sensitivity analyses, and reliable-change reporting are necessary.

4. STAC-R AND COGNITIVE SCAFFOLDING

STAC-R conceptualizes cognitive aging as the joint result of neural challenge, structural decline, compensatory recruitment, and life-course enrichment [16]. In this trial, it served only as an interpretive

account of behavioral adaptation. No neural outcome measured compensatory recruitment, connectivity, or structural change; consequently, improved test performance could reflect strategy use, motivation, familiarity, or task-specific learning rather than altered neural scaffolding.

5. AUDITORY–MOTOR ENTRAINMENT

The auditory–motor component draws on accounts in which beat perception involves predictive interaction between auditory and motor planning systems [17]. Active synchronization, rhythmic anticipation, error correction, and coordinated ensemble performance may train temporal prediction and processing efficiency. Auditory–motor synchrony was therefore specified as a music-linked process, not as a proven mediator of global cognition.

6. SOCIAL IDENTITY AND CREATIVE SELF-EFFICACY

The social-identity perspective proposes that meaningful group membership can provide belonging, support, efficacy, and shared purpose [18]. Intervention research shows that deliberately strengthening social group membership can improve mental health when isolation is a central problem [19]. Creative self-efficacy describes confidence in producing or contributing to creative outcomes; the study treated it as a modifiable belief that might support persistence, experimentation, and positive self-appraisal rather than as artistic talent [20].

7. INTEGRATED MODEL AND BOUNDARY CONDITIONS

The integrated model distinguished four levels: randomized exposures (music and performance factors); directly measured process indicators (synchrony, perceived cognitive challenge, emotion-regulation practice, creative self-efficacy, and social identification); theoretical interpretations (auditory–motor prediction, compensatory behavioral scaffolding, and group-based identity); and observed outcomes. Age, education, baseline cognition, hearing, mobility, distress, prior arts experience, attendance, group climate, facilitator fidelity, digital familiarity, and co-intervention use were boundary conditions. Dashed pathways represented exploratory associations, not verified mechanisms.

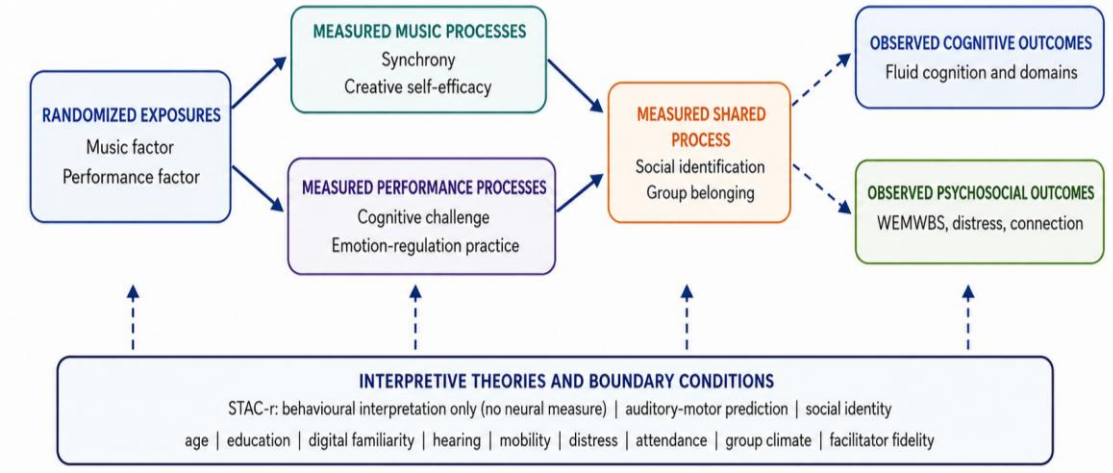


FIGURE 1. Integrated STAC-r–Auditory-Motor–Social Identity Framework. Randomized exposures, measured processes, interpretive theories, and observed outcomes are distinguished; dashed pathways are exploratory.

8. ACTIVE COMPARATORS AND FACTORIAL ASSUMPTIONS

Active comparators are essential when adult attention, structured discussion, social contact, routine, and expectancy are credible alternative explanations. They usually reduce effects relative to passive controls but

strengthen interpretation. Common practice effects are partly controlled by randomization, although unequal digital familiarity, motivation, and attrition may still bias change.

Factorial trials efficiently estimate average component effects when factors can be delivered alone and together. Interpretation depends on explicit main-effect estimands, equal contact time, separation of factor content, and reporting the interaction. A non-significant interaction does not prove additivity or equivalence when the trial is powered for main effects. Arm-specific means and confidence intervals are therefore necessary.

9. RESEARCH GAP AND ORIGINAL CONTRIBUTION

The central evidence gap is the absence of efficient active-controlled designs that separate music and performance under equal contact time while addressing practice effects, facilitator structure, safety, and durability. The study's originality is its factorial estimation of main effects for music and performance, with a secondary interaction estimate, in adults with clinically adjudicated SCD and preserved objective functioning. It does not test dementia prevention, neural scaffolding directly, or the independent effects of theatre and movement.

IV. METHODOLOGY

1. STUDY DESIGN AND REPORTING STANDARDS

A multicentre, assessor-blinded 2×2 factorial randomized controlled trial evaluated the presence or absence of music and performance. The four conditions were music only, performance only, combined music and performance, and social-learning control. Reporting followed CONSORT 2010 [21] and the factorial-trial extension [22]. The protocol was structured using SPIRIT 2013 [23] and its factorial extension [24]. We reported interventions using TIDieR [25], and we interpreted the design and process using the 2021 MRC complex-intervention framework [26].

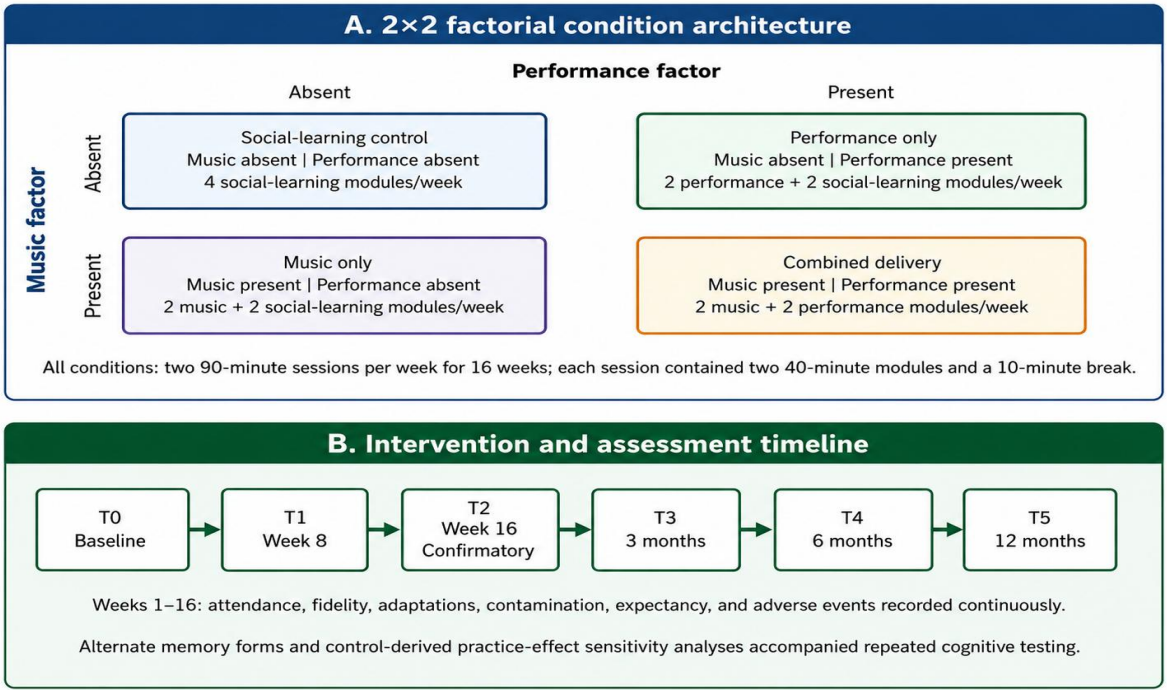


FIGURE 2. Two-panel factorial condition architecture and assessment schedule. Week 16 was the confirmatory endpoint; all conditions received equal participant contact time.

2. SETTINGS, PARTICIPANTS, AND ELIGIBILITY

Eight community aging and university-affiliated centers recruited adults aged 60–75 years. SCD eligibility required a self-reported decline lasting at least six months, associated concern, preserved independence in basic and instrumental activities, MoCA 25–30, and neuropsychological performance that did not meet mild-cognitive-impairment criteria. A domain was considered objectively impaired when two tests were at least 1.0 SD below age- and education-adjusted norms or one test was at least 1.5 SD below norms with corroborating functional concern. Two clinicians independently adjudicated eligibility based on history, informant information (when available), MoCA, memory, executive, language, and functional data; they resolved disagreements by consensus. Exclusions included dementia, major neurological disease, unstable psychiatric illness, substance dependence, severe uncorrected sensory loss, medical contraindication to moderate movement, or intensive arts/cognitive training during the previous six months.

Screening included medical history, pure-tone or validated hearing screening, corrected vision, MoCA, a brief neuropsychological battery, Amsterdam IADL interview, depressive-symptom review, and informant confirmation when available. Adults meeting mild cognitive impairment or dementia criteria were referred for clinical assessment and were not randomized. Written informed consent preceded baseline testing.

3. RANDOMIZATION, ALLOCATION, AND MASKING

After baseline assessment and clinical adjudication, an independent statistician randomized participants 1:1:1:1 using variable blocks stratified by center, age group, sex, and baseline Fluid Cognition score. A secure central system released allocation. Participants and facilitators could not be blinded because program content was visible. Cognitive assessors, outcome adjudicators, data abstractors, and the primary statistician remained masked to coded labels until the confirmatory models were locked. A masking audit showed correct guesses in 53% of sampled encounters. Participants were instructed not to disclose activities; assessments were conducted outside program rooms; and any unmasking incident was recorded before outcome scoring.

4. SAMPLE-SIZE DETERMINATION

The target was 480 participants. A conservative calculation treated the four week-16 confirmatory tests—two factors across two co-primary outcomes—as one family and used alpha=.0125 to approximate Holm control at familywise .05. Assuming repeated-measure correlation .60, delivery-group ICC .03, 18% attrition by 12 months, and baseline adjustment, the design provided approximately 80% power for a standardized main effect of .24. It was not powered for small interactions, facilitator heterogeneity, mediation, centre-specific effects, or subgroup comparisons; interaction and mechanism estimates were therefore secondary and confidence-interval based.

5. INTERVENTION CONDITIONS

All conditions received two 90-minute group sessions weekly for 16 weeks. Each session contained two 40-minute modules separated by a 10-minute break, yielding four modules and 180 participant-minutes per week. Groups contained 10–12 participants and were formed within the center after randomization using availability blocks that did not alter allocation. Progression was standardized in four phases: orientation and accessible entry (weeks 1–3), graded skill acquisition (weeks 4–7), divided attention and cooperative complexity (weeks 8–12), and consolidation with participant-led choices (weeks 13–16). We matched contact time, break structure, group size, expectancy scripts, facilitator attention, and social interaction.

Table 1. Factorial intervention conditions, exact dose, facilitator allocation, progression, and adaptations.

Domain	Social-learning control	Music only	Performance only	Combined delivery
Exact weekly dose	4 social-learning modules; 180 min/week	2 music + 2 social-learning modules; 180 min/week	2 performance + 2 social-learning modules; 180 min/week	2 music + 2 performance modules; 180 min/week
Factor content	No structured music, theatre, or choreographed movement	Voice, percussion, keyboard, ensemble rehearsal, improvisation	Theatre games, role work, narrative rehearsal, adaptable movement	Full music and full performance curricula
Core functions	Collaborative learning and cognitive discussion	Auditory–motor timing, memory, error monitoring, ensemble coordination	Embodied sequencing, role flexibility, cue response, cooperative improvisation	Both sets of core functions within equal total time
Facilitator allocation	Social-learning facilitator	Music facilitator + social-learning facilitator	Performance facilitator + social-learning facilitator	Music facilitator + performance facilitator
Progression	Orientation → graded discussion → collaborative projects → consolidation	Orientation → graded timing/sequence → divided attention → participant-led ensemble	Orientation → cue/role skills → complex enactment → participant-led scenes	Parallel progression of both factors without extra dose
Permitted adaptations	Pacing, print size, hearing support, breaks	Lower sound, amplification, simplified parts, visual cues, seated access	Seated movement, reduced range, rehearsal privacy, non-performance roles	All factor-specific adaptations with fatigue monitoring
Prohibited content	Active music-making, dramatic enactment, choreographed movement	Theatre rehearsal and performance movement	Structured music-making and rhythm training	Additional social-learning module
Fidelity threshold	≥85% control elements; no prohibited factor content	≥85% music elements; no performance content	≥85% performance elements; no music content	≥85% of each factor

6. MUSIC-MAKING FACTOR

Music modules used active voice, hand percussion, accessible keyboards, melodic and rhythmic imitation, call-and-response, group arrangement, and structured improvisation. Difficulty increased through longer sequences, divided attention, tempo change, harmonic choice, and rotating leadership. No prior musical experience was required. Facilitators prioritized accurate engagement and shared timing over performance quality.

7. PERFORMANCE FACTOR

Performance modules integrated theatre and movement around role adoption, narrative recall, cue response, spatial patterning, gesture, partner mirroring, improvisation, and short ensemble scenes. Movement was adaptable to seated or standing participation. Public performance was optional; the primary task was rehearsal and cooperative enactment, not presentation to an audience.

8. SOCIAL-LEARNING COMPARATOR

Social-learning modules included facilitated discussion, local history, health literacy, problem-solving quizzes, documentary excerpts, memory strategies, and collaborative planning. They excluded active music-making, dramatic role enactment, choreographed movement, and rhythm training. The comparator was designed to match contact, attention, social interaction, cognitive discussion, and expectancy without reproducing the hypothesized active processes.

9. FACILITATOR TRAINING, FIDELITY, AND SAFETY

Sixteen music facilitators, sixteen performance facilitators, and twelve social-learning facilitators completed protocol, safeguarding, accessibility, adverse-event, and fidelity training. The same music facilitators delivered music modules in music-only and combined groups; the same performance facilitators delivered performance modules in performance-only and combined groups. Social-learning facilitators delivered control and single-factor modules. Participants could therefore be exposed to more than one facilitator type, and exposure minutes were recorded for multiple-membership modeling. No facilitator worked outside professional competence. Independent observers rated 25% of sessions, 20% were double-coded, and facilitator rotation or substitution was logged.

Adherence was defined prospectively as attendance at at least 24 of 32 sessions and receipt of at least 75% of assigned factor modules. Permitted adaptations included seated movement, reduced range, lower sound level, amplification, visual cueing, simplified instrument parts, pacing changes, rehearsal privacy, and non-performance roles. Prohibited cross-factor content was audited using module checklists and independent observation. Contamination was classified as minor when brief and incidental, and substantial when an unassigned factor occupied at least 20% of a module or occurred in three or more sessions.

Adverse events included musculoskeletal discomfort, falls, dizziness, vocal strain, hearing discomfort, emotional activation, interpersonal conflict, fatigue, and frustration. Serious events were reviewed by an independent clinician. Participants could sit out, reduce movement range, use hearing support, lower sound intensity, or discontinue an exercise without leaving outcome follow-up.

10. OUTCOMES AND MEASUREMENT SCHEDULE

The first co-primary outcome was the NIH Toolbox Fluid Cognition Composite. The primary score was the age-unadjusted standard score, with baseline age, sex, education, and digital familiarity included as covariates; fully demographically corrected scores were examined in sensitivity analysis. Alternate forms were used where available for episodic-memory measures, and task order was counterbalanced. We prespecified a five-point increase as a pragmatic reliable-improvement threshold based on test-retest precision, not a clinical cut-off [27]. The second co-primary outcome was the 14-item WEMWBS; a three-point increase was used as a pragmatic meaningful-change threshold and reported alongside continuous effects [28].

Secondary outcomes were NIH Toolbox component scores, Trail Making switching, alternate-form Rey Auditory Verbal Learning delayed recall, K10 psychological distress, Lubben Social Network Scale-6, Amsterdam IADL short form, PROMIS Sleep Disturbance 8b, EQ-5D-5L index and VAS, attendance, fidelity, adaptations, contamination, and adverse events. The Amsterdam IADL short form is sensitive to early functional decline [29], and PROMIS Sleep Disturbance has acceptable performance in older adults [30]. Practice-effect sensitivity analyses subtracted the control group's mean retest change from each cognitive outcome and repeated models using demographically corrected NIH scores.

https://doi.org/10.48161/aaj.v3n2a97

Table 2. Instruments, scoring, informants, assessment occasions, and analytical roles. T0, baseline; W, week; m, months.

Domain	Instrument / version	Score and informant	Assessment occasions	Role	Measurement note
Fluid cognition	NIH Toolbox Fluid Cognition Composite	Age-unadjusted standard score; assessor administered	T0, W8, W16, 3m, 6m, 12m	Co-primary	Fully corrected and practice-adjusted sensitivity analyses
Mental well-being	WEMWBS, 14-item	0–70 participant report	T0, W8, W16, 3m, 6m, 12m	Co-primary	Three-point pragmatic responder threshold
Executive function	NIH Flanker/DCCS; Trail switching	Standard scores / completion contrast	T0, W8, W16, 3m, 6m, 12m	Secondary	Assessor masked
Episodic memory	NIH Picture Sequence; alternate-form RAVLT delayed	Standard / raw delayed recall	T0, W16, 3m, 6m, 12m	Secondary	Alternate forms counterbalanced
Processing speed	NIH Pattern Comparison	Standard score	T0, W8, W16, 3m, 6m, 12m	Secondary	Digital familiarity adjusted
Psychological distress	Kessler-10	10–50 participant report; lower is better	T0, W8, W16, 3m, 6m, 12m	Secondary	Psychosocial FDR family
Social connectedness	Lubben Social Network Scale-6	0–30 participant report	T0, W8, W16, 3m, 6m, 12m	Secondary	Psychosocial FDR family
Everyday function	Amsterdam IADL short form	Participant/informant standardized score	T0, W16, 3m, 6m, 12m	Secondary	Informant where available
Sleep	PROMIS Sleep Disturbance 8b	T-score; lower better	T0, W16, 3m, 6m, 12m	Secondary	No W8 assessment
Quality of life	EQ-5D-5L index and VAS	Participant report	T0, W16, 3m, 6m, 12m	Secondary	No W8 assessment
Processes	Synchrony; challenge; regulation; creative self-efficacy; social identification	Objective tapping + study scales	T0, W4, W8, W16	Exploratory	Internal validation only
Implementation / safety	Attendance, dose, fidelity, adaptations, contamination, adverse events	Logs and independent observation	Continuous during weeks 1–16	Implementation	Exposure-hour denominators

11. MECHANISM MEASURES

Auditory–motor synchrony was measured with a sensor-timed paced-tapping task using circular phase consistency and test–retest ICC=.82. Perceived cognitive challenge used a six-item study-specific scale (omega=.84); emotion-regulation practice used a six-item reappraisal and regulation index (omega=.81); creative self-efficacy used three items adapted from the framework cited in [20] (omega=.86); and social identification used a four-item group-identification measure (omega=.88). A prespecified confirmatory-factor model showed acceptable internal fit (CFI=.94, TLI=.92, RMSEA=.052), and partial longitudinal scalar invariance was retained after freeing two thresholds. These findings support exploratory internal use only; they do not establish independent validation, criterion validity, or causal-mechanism status.

12. QUANTITATIVE ANALYSIS

Analyses followed intention-to-treat principles. For participant i in delivery group j at centre k and occasion t, the primary model was $Y_{ijkt} = \beta0 + \beta1\text{Music} + \beta2\text{Performance} + \beta3\text{Music}\times\text{Performance} + \beta4\text{Time} + \beta5\text{Music}\times\text{Time} + \beta6\text{Performance}\times\text{Time}$ + baseline and stratification covariates + random participant and group effects + exposure-weighted music-, performance-, and social-facilitator effects + residual error. Facilitator weights were proportional to attended module minutes, producing a multiple-membership structure for participants receiving more than one facilitator type. Centre was included as a fixed effect; repeated residuals used an unstructured covariance through week 16 and autoregressive decay across follow-up. Kenward–Roger degrees of freedom addressed finite groups and facilitators. Arm-specific adjusted means accompanied factorial main effects and interactions.

The four week-16 factorial main-effect tests were controlled as one family using Holm's step-down procedure at familywise alpha=.05. Interactions were secondary and interpreted using estimates and 95% confidence intervals rather than significance as proof of additivity. Secondary cognitive outcomes formed one false-discovery-rate family, and psychosocial/functional outcomes formed another, with separate q values for music and performance effects. Hedges' g used the pooled baseline SD. Prespecified pragmatic responder analyses reported five-point Fluid Cognition and three-point WEMWBS improvements. Group and facilitator ICCs, model degrees of freedom, residual fit, convergence, and influential centres were reported.

Missing outcomes were handled using 50 multilevel chained-equation imputations incorporating allocation, centre, delivery group, exposure-weighted facilitators, all outcome occasions, baseline cognition, distress, digital familiarity, hearing, mobility, attendance, expectations, co-interventions, adverse events, and reasons for missingness [31]. Sensitivity analyses included complete cases, no imputation, per-protocol exposure, alternative covariance structures, fully corrected NIH scores, control-derived practice adjustment, exclusion of substantial contamination, physical-activity and cognitive-training adjustment, inverse-probability retention weighting, pattern-mixture departures from missing at random, and leave-one-center-out models.

Factorial analyses can be biased when plausible interactions are ignored; therefore, arm-specific means and interaction intervals accompanied main effects [32]. The study could detect only moderate interactions with useful power, so null interaction estimates did not demonstrate equivalence. Longitudinal process models used week-4 and week-8 measures and week-16 outcomes with group-robust uncertainty. Because process measures were non-randomized and mediator–outcome confounding could not be eliminated, results were described as indirect associations rather than causal mediation [33].

13. ETHICS, PRIVACY, AND DATA GOVERNANCE

The coordinating university research ethics committee approved the study before recruitment, and each centre completed governance review. Participants gave written informed consent covering cognitive testing, questionnaires, session observation, adverse-event monitoring, and controlled data sharing. Participation was voluntary, withdrawal did not affect services, and participants could discontinue any activity while remaining in follow-up.

We stored identifying information separately from coded research data. Electronic records were encrypted on role-restricted institutional systems with access logs, secure transfer, and prespecified retention

periods. We stored paper forms in locked cabinets. Reports used aggregated results only. A data-monitoring group reviewed recruitment, missingness, fidelity, and adverse events without access to unmasked comparative outcomes.

V. RESULTS

1. PARTICIPANT FLOW AND BASELINE CHARACTERISTICS

Across eight centres, 638 adults were screened. Exclusion reasons were objective impairment or mild cognitive impairment (n=54), dementia or major neurological disorder (n=19), medical, sensory, or movement contraindications (n=31), and declining participation (n=54). The remaining 480 completed baseline assessment and were randomized equally. We observed week-16 outcomes for 446 participants (92.9%), three-month outcomes for 426 (88.8%), six-month outcomes for 407 (84.8%), and 12-month outcomes for 388 (80.8%). Missingness was associated with baseline distress, transport difficulty, and lower attendance but not independently with either randomized factor or baseline co-primary outcome. All randomized participants contributed to imputed models.

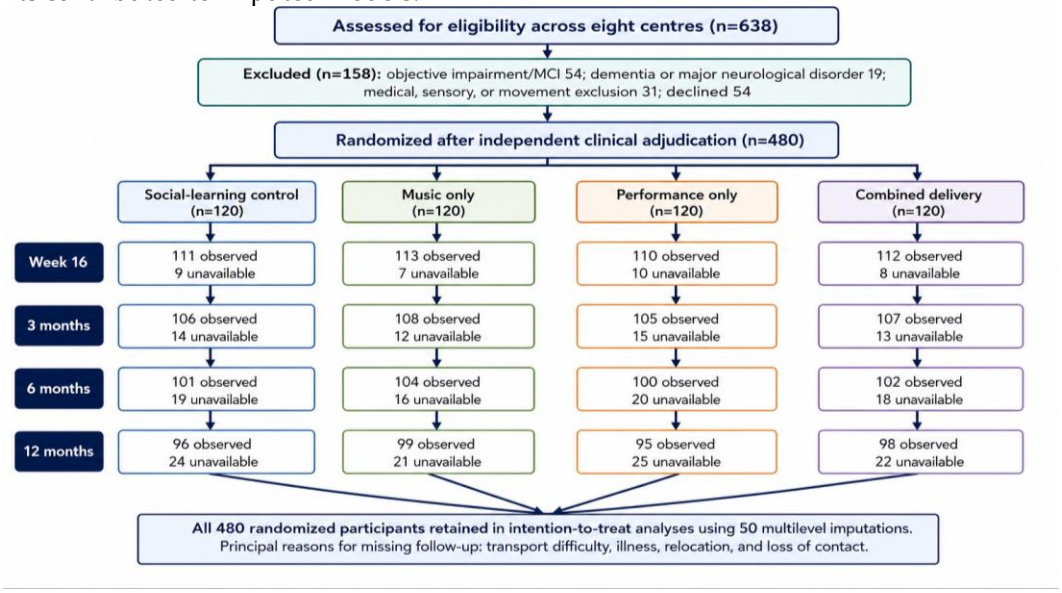


FIGURE 3. Participant flow through the factorial trial. Eligibility exclusions and principal reasons for missing follow-up are shown.

Table 3. Baseline participant characteristics. SMD, standardized mean difference.

Characteristic	Control (n=120)	Music only (n=120)	Performance only (n=120)	Combined (n=120)	Max. SMD
Age, years, mean (SD)	67.3 (4.2)	67.1 (4.1)	67.5 (4.0)	67.2 (4.3)	0.10
Women, n (%)	72 (60.0)	70 (58.3)	73 (60.8)	71 (59.2)	0.05
University education, n (%)	48 (40.0)	50 (41.7)	47 (39.2)	49 (40.8)	0.05
Ethnic-minority background, n (%)	31 (25.8)	29 (24.2)	32 (26.7)	30 (25.0)	0.06

Characteristic	Control (n=120)	Music only (n=120)	Performance only (n=120)	Combined (n=120)	Max. SMD
Lower neighbourhood-income tertile, n (%)	36 (30.0)	35 (29.2)	37 (30.8)	36 (30.0)	0.04
Centre distribution, range per centre	13–17	13–17	13–17	13–17	0.08
Subjective decline duration, years	2.8 (1.5)	2.7 (1.4)	2.9 (1.6)	2.8 (1.5)	0.13
MoCA, mean (SD)	26.4 (1.5)	26.5 (1.4)	26.3 (1.6)	26.4 (1.5)	0.13
NIH Fluid Cognition	94.8 (11.8)	94.7 (11.7)	94.9 (11.6)	94.8 (11.9)	0.02
WEMWBS	43.1 (10.8)	43.2 (10.6)	43.0 (10.9)	43.3 (10.7)	0.03
Amsterdam IADL standardized score	58.7 (7.2)	58.9 (7.0)	58.5 (7.4)	58.8 (7.1)	0.06
Low digital familiarity, n (%)	27 (22.5)	28 (23.3)	26 (21.7)	27 (22.5)	0.04
Hearing aid use, n (%)	24 (20.0)	22 (18.3)	25 (20.8)	23 (19.2)	0.06
Mild hearing loss on screening, n (%)	39 (32.5)	37 (30.8)	40 (33.3)	38 (31.7)	0.06
Mobility limitation, n (%)	19 (15.8)	18 (15.0)	20 (16.7)	18 (15.0)	0.05
Prior organized music, n (%)	31 (25.8)	30 (25.0)	32 (26.7)	31 (25.8)	0.04
Prior theatre/dance, n (%)	22 (18.3)	23 (19.2)	21 (17.5)	22 (18.3)	0.04

2. *EXPOSURE, FIDELITY, ACCEPTABILITY, AND SAFETY*

Median attendance was 28 of 32 sessions in control, 29 in music only, 28 in performance only, and 29 in combined delivery. Eighty-one percent of participants attended at least 24 sessions. Independent observers assessed 128 sessions per factor. Mean music fidelity was 91% in music-only and 90% in combined groups; performance fidelity was 89% and 88%, respectively. Social-learning fidelity was 90%. Inter-rater ICCs ranged from .82 to .88. Cross-factor contamination was documented in 2.1% of audited sessions and was minor.

Acceptability averaged 4.3/5 for music, 4.1/5 for performance, 4.2/5 for combined delivery, and 3.9/5 for control. Combined participants reported greater fatigue at week 8, but withdrawal did not differ materially. Forty-six intervention-related adverse events occurred: transient musculoskeletal discomfort (18), vocal or hearing fatigue (9), dizziness (5), emotional activation (6), frustration (6), and minor interpersonal conflict (2). Two non-injurious falls occurred during performance modules. No serious event was attributed to study participation.

Table 4. Attendance, fidelity, acceptability, contamination, and adverse-event exposure rates.

Indicator	Control	Music only	Performance only	Combined
Median sessions attended (IQR)	28 (25–31)	29 (26–31)	28 (24–31)	29 (25–31)
Attended ≥24 sessions	94/120 (78.3%)	99/120 (82.5%)	96/120 (80.0%)	100/120 (83.3%)
Core fidelity	90%	91%	89%	Music 90%; performance 88%
Acceptability, mean /5	3.9	4.3	4.1	4.2
Mild/moderate related events, n participants	7/120	9/120	13/120	17/120
Related-event rate per 1,000 participant-hours	1.39	1.72	2.58	3.26
Event severity: mild / moderate	6 / 1	8 / 1	10 / 3	13 / 4
Withdrawals attributed to burden	1	1	2	2
Serious intervention-related events	0	0	0	0
Substantial contamination	2 groups	1 group	1 group	1 group

3. CO-PRIMARY FACTORIAL EFFECTS

At the confirmatory week-16 endpoint, music produced a +2.6-point Fluid Cognition main effect (95% CI 1.1–4.1; g=.22; Holm-adjusted p=.003), and performance produced +2.1 points (95% CI 0.6–3.6; g=.18; p=.006). The interaction was −0.4 (95% CI −2.3–1.5); this interval did not clearly depart from additivity but allowed for a small benefit or attenuation. Adjusted means were 95.4 in control, 98.1 in music only, 97.5 in performance only, and 99.5 in combined delivery.

Music produced a +2.2-point WEMWBS main effect at week 16 (95% CI 0.8–3.6; g=.20; Holm-adjusted p=.004), and performance produced +2.8 points (95% CI 1.4–4.2; g=.25; p<.002). The interaction was +0.3 (95% CI −1.5–2.1), again too imprecise to establish additivity or equivalence. The four adjusted means were 44.0, 46.1, 46.8, and 49.0. The combined condition had the largest descriptive mean because both factor estimates were positive; synergy was not demonstrated.

At three months, music and performance cognition effects were +2.0 and +1.7 points; at six months, +1.3 and +1.0; and at 12 months, +0.7 and +0.5, with 12-month intervals crossing no difference. WEMWBS followed the same attenuation pattern. Practice-adjusted sensitivity models reduced the week-16 cognition estimates slightly to +2.4 for music and +1.9 for performance, without changing interpretation. Five-point Fluid Cognition improvement occurred in 42.1% with music present versus 31.3% with music absent (adjusted RR 1.35, 95% CI 1.06–1.72) and in 39.6% with performance present versus 33.8% absent (RR 1.17, 95% CI 0.92–1.49). Three-point WEMWBS improvement occurred in 52.5% with music present versus 43.3% absent and 55.8% with performance present versus 40.0% absent. These were pragmatic responder analyses, not clinical recovery thresholds.

Table 5. Confirmatory week-16 factorial estimates. Holm-adjusted p-values control the four primary tests; interaction intervals remain secondary.

Co-primary outcome	Adjusted arm means: control | music | performance | combined	Music main effect (95% CI); Holm p	Performance main effect (95% CI); Holm p	Interaction (95% CI)	Group / facilitator ICC	df
Fluid Cognition, week 16	95.4 | 98.1 | 97.5 | 99.5	+2.6 (1.1, 4.1); .003	+2.1 (0.6, 3.6); .006	−0.4 (−2.3, 1.5)	0.028 / 0.013	42.6
WEMWBS, week 16	44.0 | 46.1 | 46.8 | 49.0	+2.2 (0.8, 3.6); .004	+2.8 (1.4, 4.2); <.002	+0.3 (−1.5, 2.1)	0.024 / 0.010	40.9

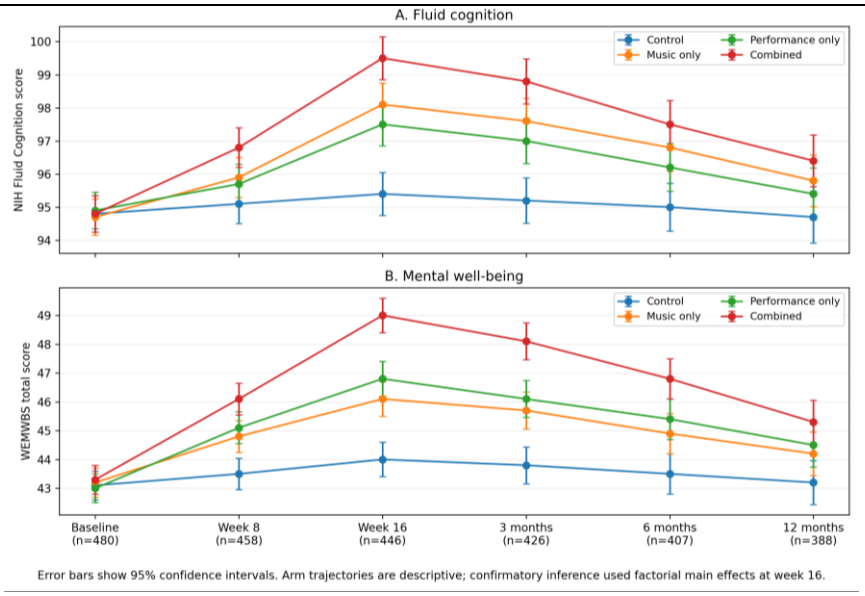


FIGURE 4. Descriptive randomized-condition trajectories with 95% confidence intervals. Confirmatory inference used factorial main effects at week 16.

4. SECONDARY COGNITIVE, PSYCHOLOGICAL, AND FUNCTIONAL OUTCOMES

Secondary effects were selective. Music had the clearest estimates for processing speed and episodic memory; performance had the clearest estimates for executive switching, psychological distress, and social connectedness. Separate q values showed that statistical reliability for one factor did not establish specificity relative to the other. After multiplicity control, everyday function, sleep, and quality of life remained small or uncertain.

Secondary interaction intervals were centred near zero but were wide enough to permit small departures from additivity. A negative week-8 fatigue interaction suggested greater early burden in combined delivery, which attenuated by week 16. Descriptive combined-arm gains were not interpreted as categorical superiority.

Table 6. Week-16 secondary factorial effects with separate false-discovery-rate q values for each factor.

Week-16 outcome	Music effect (95% CI)	Music q	Performance effect (95% CI)	Performance q	Interpretation
Processing speed	+2.4 (0.9, 3.9)	.012	+0.8 (−0.7, 2.3)	.380	Clearer music main effect
Episodic memory	+1.9 (0.4, 3.4)	.036	+1.0 (−0.5, 2.5)	.220	Music-favouring pattern
Working memory	+1.7 (0.2, 3.2)	.064	+1.2 (−0.3, 2.7)	.180	Small and uncertain after FDR
Executive switching	+0.9 (−0.6, 2.4)	.310	+2.2 (0.7, 3.7)	.024	Clearer performance main effect
K10 psychological distress	−1.2 (−2.4, 0.0)	.082	−1.8 (−3.0, −0.6)	.019	Performance-favouring pattern
Social connectedness	+1.4 (0.1, 2.7)	.048	+2.6 (1.3, 3.9)	.006	Performance estimate larger
Amsterdam IADL	+0.8 (−0.3, 1.9)	.190	+1.1 (0.0, 2.2)	.120	Uncertain small effects
PROMIS sleep disturbance	−0.7 (−1.8, 0.4)	.260	−0.8 (−1.9, 0.3)	.240	No reliable effect
EQ-5D-5L VAS	+1.5 (0.1, 2.9)	.071	+1.7 (0.3, 3.1)	.058	Small; FDR uncertain

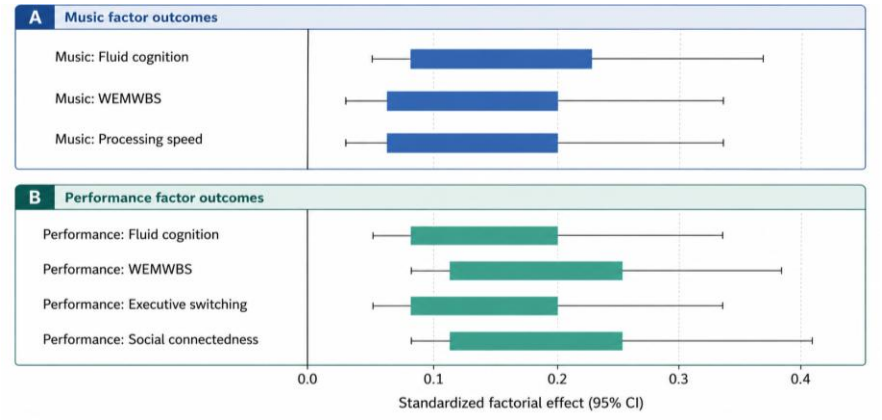


FIGURE 5. Neutral-color Forest plot of principal week-16 factorial effects. Positive values favor assignment to the relevant factor.

5. EXPLORATORY MECHANISMS AND MODERATORS

By week 4, music assignment increased auditory–motor synchrony and creative self-efficacy, while performance increased perceived cognitive challenge, emotion-regulation practice, and social identification. Temporally ordered models identified small indirect associations in theoretically expected directions. Estimates were correlated and changed after attendance, expectancy, group climate, and facilitator-fidelity adjustment.

No process estimate was treated as a confirmed mechanism, and no moderator remained reliable after multiplicity control. Higher baseline distress and lower digital familiarity showed suggestive variation, but the trial was not powered to define responsive subgroups. Complete mediation, moderator, and interaction results are reported in the Supplementary Material.

VI. DISCUSSION

The confirmatory analyses showed small week-16 main effects of both music and performance on fluid cognition and mental well-being after Holm control of the four primary tests. Music had a slightly larger cognitive estimate, while performance had a slightly larger well-being estimate. The interaction intervals did not show clear departure from additivity, but the trial lacked power to exclude small synergistic or attenuating effects. Therefore, the combined arm cannot be described as synergistic or equivalent to the sum of its components.

The cognitive findings are consistent with the small average effects reported in music and broader arts reviews [3]–[6]. The active comparator matched social contact, adult attention, predictable routine, and cognitive discussion, so the estimates represent added factor content rather than the full effect of joining a group programme. The active comparator may partly account for the smaller effect estimates than those observed in passive-control studies. Practice-adjusted and fully corrected NIH analyses were similar, although repeated testing cannot be ruled out as a source of influence.

Music showed clearer estimates for processing speed and episodic memory. Auditory prediction, timed motor coordination, repeated retrieval, and error monitoring offer plausible behavioural explanations. STAC-r remains interpretive because no neural measure assessed compensatory recruitment. Changes could also reflect strategy acquisition, engagement, or task familiarity. Synchrony and cognitive-challenge associations were small and confounded by attendance and facilitator fidelity.

Performance showed clearer estimates for executive switching, distress, and social connectedness. Theatre and movement required role change, response inhibition, perspective shifting, narrative memory, spatial adaptation, and interpersonal coordination. Social identification and emotion-regulation practice were compatible with the well-being effect, but neither process was randomized. Because theatre and movement were bundled, the trial cannot identify which performance component contributed to change.

The responder analyses provided practical context but did not establish clinical recovery. More participants assigned to music achieved a five-point increase in Fluid Cognition, and more participants assigned to each factor achieved a three-point increase in WEMWBS. These thresholds were pragmatic and based on measurement precision or prior population interpretation, not disease-specific minimal clinically important differences. Everyday function, sleep, and broad quality of life changed little, limiting claims of generalized health improvement.

Effects weakened after organized delivery and were uncertain at 12 months. The study did not determine whether continued practice or booster sessions would maintain effects. SCD is heterogeneous, and many participants will not develop objective impairment. The findings therefore do not address dementia incidence, neurodegeneration, or durable cognitive protection.

Implementation was feasible but resource-dependent. Attendance and fidelity were high with trained facilitators, a protected space, accessible instruments, adaptable movement, and a structured progression. Mild and moderate burdens—including musculoskeletal discomfort, sound fatigue, dizziness, emotional activation, and performance anxiety—were more frequent in the combined arm. Exposure-hour rates and withdrawals show why safety cannot be summarized only by the absence of serious events.

Methodologically, the factorial estimands were more informative than treating four arms as unrelated comparisons. Multiple-membership models represented exposure to different facilitator types, and group and facilitator ICCs were small but non-zero. Nonetheless, facilitator expertise remained partly linked to factor content, interaction power was limited, and participants and facilitators were unblinded. These conditions permit expectancy and performance effects, particularly for WEMWBS.

Additional limitations include clinically heterogeneous SCD, computerized assessment influenced by digital familiarity, 19.2% attrition at 12 months, and possible missing-not-at-random mechanisms despite imputation and pattern-mixture analyses. Process scales showed acceptable internal properties but lacked independent validation. Center-specific and subgroup estimates were imprecise. The trial excluded dementia and objectively impaired mild cognitive impairment, limiting generalization to clinical cognitive disorders.

The findings support offering music and performance as preference-sensitive community options when the goal is short-term cognitive stimulation, well-being, and social participation. Neither factor was globally

superior. Combined delivery may offer variety but requires fatigue monitoring and does not justify a synergy claim. Cultural adaptation should preserve active sound production and temporal coordination for music, and role flexibility, embodied sequencing, and cooperative improvisation for performance, while maintaining equivalent dose, accessibility, and fidelity.

VII. CONCLUSION

In adults aged 60–75 years with clinically adjudicated subjective cognitive decline but no dementia or mild cognitive impairment, 16 weeks of active group music-making and theatre- and movement-based performance each produced small improvements in fluid cognition and mental well-being relative to an attention- and time-matched social-learning programme. Holm-corrected week-16 results supported both factorial main effects, while interaction confidence intervals did not establish synergy, antagonism, additivity, or equivalence. Music showed clearer estimates for processing speed and episodic memory, whereas performance showed clearer estimates for executive switching, psychological distress, and social connectedness. Effects attenuated after delivery and were uncertain at 12 months. Pragmatic responder analyses indicated that some participants experienced meaningful short-term change, but everyday function, sleep, and broad quality of life showed limited differences. Exploratory associations involving synchrony, cognitive challenge, emotion regulation, creative self-efficacy, and social identification remained confounded and did not verify causal mechanisms. The programmes were feasible under trained facilitators and accessibility adaptations, although musculoskeletal discomfort, fatigue, sound sensitivity, dizziness, emotional activation, and performance anxiety required monitoring. The findings support preference-sensitive music and performance programs as optional short-term cognitive and psychosocial resources. They do not demonstrate dementia prevention, durable cognitive protection, component-specific theatre or movement effects, or synergy from combined delivery.

VIII. DATA AVAILABILITY AND SUPPLEMENTARY MATERIAL

De-identified quantitative data, data dictionaries, analysis code, fidelity tools, and adverse-event coding materials are governed through controlled access subject to methodological review, ethics approval, and a data-use agreement.

The supplement contains the protocol synopsis, statistical analysis plan, factorial estimands, multiple-membership model equations, multiplicity procedure, interaction-power statement, imputation diagnostics, complete sensitivity, process and moderator estimates, facilitator structure, fidelity tools, reporting checklists, and adverse-event definitions.

Author Contributions

The author conducted the conceptualization, methodology, data analysis, investigation, writing, review, editing, and final approval of the manuscript.

Funding

This research received no external funding.

Data Availability

The dataset will be available from the author upon reasonable request.

Conflicts of Interest

The author declares no conflict of interest.

REFERENCES

[1]. F. Jessen, R. E. Amariglio, M. van Boxtel, et al., "A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease," Alzheimer's & Dementia, vol. 10, no. 6, pp. 844–852, 2014, doi: 10.1016/j.jalz.2014.01.001.

[2]. G. Livingston, J. Huntley, A. Sommerlad, et al., "Dementia prevention, intervention, and care: 2020 report of the Lancet Commission," The Lancet, vol. 396, no. 10248, pp. 413–446, 2020, doi: 10.1016/S0140-6736(20)30367-6.
[3]. Z. H. Fong, S. H. Tan, R. Mahendran, E. H. Kua, and T. T. Chee, "Arts-based interventions to improve cognition in older persons with mild cognitive impairment: A systematic review of randomized controlled trials," Aging & Mental Health, vol. 25, no. 9, pp. 1605–1617, 2021, doi: 10.1080/13607863.2020.1786802.
[4]. J. L. Dorris, S. Neely, L. Terhorst, H. M. VonVille, and J. Rodakowski, "Effects of music participation for mild cognitive impairment and dementia: A systematic review and meta-analysis," Journal of the American Geriatrics Society, vol. 69, no. 9, pp. 2659–2667, 2021, doi: 10.1111/jgs.17208.
[5]. M. Bleibel, A. El Cheikh, N. S. Sadier, and L. Abou-Abbas, "The effect of music therapy on cognitive functions in patients with Alzheimer's disease: A systematic review of randomized controlled trials," Alzheimer's Research & Therapy, vol. 15, art. 65, 2023, doi: 10.1186/s13195-023-01214-9.
[6]. G. Ma and X. Ma, "Music intervention for older adults: Evidence map of systematic reviews," Medicine, vol. 102, no. 48, art. e36016, 2023, doi: 10.1097/MD.0000000000036016.
[7]. I. C. Galinha, M. Á. García-Martín, and M. L. Lima, "Sing4Health: Randomised controlled trial of the effects of a singing group program on the subjective and social well-being of older adults," Applied Psychology: Health and Well-Being, vol. 14, no. 1, pp. 176–195, 2022, doi: 10.1111/aphw.12297.
[8]. J. K. Johnson, A. L. Stewart, M. Acree, et al., "A community choir intervention to promote well-being among diverse older adults: Results from the Community of Voices trial," The Journals of Gerontology: Series B, vol. 75, no. 3, pp. 549–559, 2020, doi: 10.1093/geronb/gby132.
[9]. N. Daykin, A. Mansfield, C. Meads, et al., "What works for wellbeing? A systematic review of wellbeing outcomes for music and singing in adults," Perspectives in Public Health, vol. 138, no. 1, pp. 39–46, 2018, doi: 10.1177/1757913917740391.
[10]. D. Predovan, A. Julien, A. Esmail, and L. Bherer, "Effects of dancing on cognition in healthy older adults: A systematic review," Journal of Cognitive Enhancement, vol. 3, no. 2, pp. 161–167, 2019, doi: 10.1007/s41465-018-0103-2.
[11]. X. Meng, G. Li, Y. Jia, et al., "Effects of dance intervention on global cognition, executive function and memory of older adults: A meta-analysis and systematic review," Aging Clinical and Experimental Research, vol. 32, no. 1, pp. 7–19, 2020, doi: 10.1007/s40520-019-01159-w.
[12]. A. Kodama, N. Watanabe, H. Ozawa, et al., "Senior theater projects: Enhancing physical health and reducing depression in older adults," International Journal of Environmental Research and Public Health, vol. 21, no. 10, art. 1289, 2024, doi: 10.3390/ijerph21101289.
[13]. C. Liu, M. Su, Y. Jiao, Y. Ji, and S. Zhu, "Effects of dance interventions on cognition, psycho-behavioral symptoms, motor functions, and quality of life in older adult patients with mild cognitive impairment: A meta-analysis and systematic review," Frontiers in Aging Neuroscience, vol. 13, art. 706609, 2021, doi: 10.3389/fnagi.2021.706609.
[14]. J. S. Y. Chan, J. Wu, K. Deng, and J. H. Yan, "The effectiveness of dance interventions on cognition in patients with mild cognitive impairment: A meta-analysis of randomized controlled trials," Neuroscience & Biobehavioral Reviews, vol. 118, pp. 80–88, 2020, doi: 10.1016/j.neubiorev.2020.07.017.
[15]. B. N. Sprague, S. A. Freed, C. E. Webb, C. B. Phillips, J. Hyun, and L. A. Ross, "The impact of behavioral interventions on cognitive function in healthy older adults: A systematic review," Ageing Research Reviews, vol. 52, pp. 32–52, 2019, doi: 10.1016/j.arr.2019.04.002.
[16]. P. A. Reuter-Lorenz and D. C. Park, "How does it STAC up? Revisiting the scaffolding theory of aging and cognition," Neuropsychology Review, vol. 24, no. 3, pp. 355–370, 2014, doi: 10.1007/s11065-014-9270-9.
[17]. A. D. Patel and J. R. Iversen, "The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction hypothesis," Frontiers in Systems Neuroscience, vol. 8, art. 57, 2014, doi: 10.3389/fnsys.2014.00057.
[18]. J. Jetten, S. A. Haslam, T. Cruwys, K. H. Greenaway, C. Haslam, and N. K. Steffens, "Advancing the social identity approach to health and well-being: Progressing the social cure research agenda," European Journal of Social Psychology, vol. 47, no. 7, pp. 789–802, 2017, doi: 10.1002/ejsp.2333.
[19]. T. Cruwys, C. Haslam, G. A. Dingle, J. Jetten, K. Hornsey, and S. A. Haslam, "Groups 4 Health: Evidence that a social-identity intervention that builds and strengthens social group membership improves mental health," Journal of Affective Disorders, vol. 194, pp. 188–195, 2016, doi: 10.1016/j.jad.2016.01.010.
[20]. M. Karwowski, "The dynamics of creative self-concept: Changes and reciprocal relations between creative self-efficacy and creative personal identity," Creativity Research Journal, vol. 28, no. 1, pp. 99–104, 2016, doi: 10.1080/10400419.2016.1125254.

[21]. K. F. Schulz, D. G. Altman, and D. Moher, "CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomized trials," BMJ, vol. 340, art. c332, 2010, doi: 10.1136/bmj.c332.

[22]. B. C. Kahan, E. Juszczak, E. M. Beller, et al., "Reporting of factorial randomized trials: Extension of the CONSORT 2010 Statement," JAMA, vol. 330, no. 21, pp. 2106–2114, 2023, doi: 10.1001/jama.2023.19793.

[23]. A.-W. Chan, J. M. Tetzlaff, D. G. Altman, et al., "SPIRIT 2013 statement: Defining standard protocol items for clinical trials," Annals of Internal Medicine, vol. 158, no. 3, pp. 200–207, 2013, doi: 10.7326/0003-4819-158-3-201302050-00583.

[24]. B. C. Kahan, E. Juszczak, E. M. Beller, et al., "Consensus statement for protocols of factorial randomized trials: Extension of the SPIRIT 2013 Statement," JAMA Network Open, vol. 6, no. 12, art. e2346121, 2023, doi: 10.1001/jamanetworkopen.2023.46121.

[25]. T. C. Hoffmann, P. P. Glasziou, I. Boutron, et al., "Better reporting of interventions: Template for Intervention Description and Replication (TIDieR) checklist and guide," BMJ, vol. 348, art. g1687, 2014, doi: 10.1136/bmj.g1687.

[26]. K. Skivington, L. Matthews, S. A. Simpson, et al., "A new framework for developing and evaluating complex interventions: Update of Medical Research Council guidance," BMJ, vol. 374, art. n2061, 2021, doi: 10.1136/bmj.n2061.

[27]. E. P. Scott, A. Sorrell, and A. Benitez, "Psychometric properties of the NIH Toolbox Cognition Battery in healthy older adults: Reliability, validity, and agreement with standard neuropsychological tests," Journal of the International Neuropsychological Society, vol. 25, no. 8, pp. 857–867, 2019, doi: 10.1017/S1355617719000614.

[28]. R. Tennant, L. Hiller, R. Fishwick, et al., "The Warwick–Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation," Health and Quality of Life Outcomes, vol. 5, art. 63, 2007, doi: 10.1186/1477-7525-5-63.

[29]. R. J. Jutten, C. F. W. Peeters, S. M. J. Leijdesdorff, et al., "Detecting functional decline from normal aging to dementia: Development and validation of a short version of the Amsterdam IADL Questionnaire," Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, vol. 8, pp. 26–35, 2017, doi: 10.1016/j.dadm.2017.03.002.

[30]. J. M. Hughes, Y. Song, P. K. Fung, et al., "Assessing psychometric properties of the PROMIS Sleep Disturbance Scale in older adults in independent-living and continuing care retirement communities," Sleep Health, vol. 4, no. 6, pp. 552–559, 2018, doi: 10.1016/j.sleh.2018.09.003.

[31]. I. R. White, P. Royston, and A. M. Wood, "Multiple imputation using chained equations: Issues and guidance for practice," Statistics in Medicine, vol. 30, no. 4, pp. 377–399, 2011, doi: 10.1002/sim.4067.

[32]. B. C. Kahan, "Bias in randomized factorial trials," Statistics in Medicine, vol. 32, no. 26, pp. 4540–4549, 2013, doi: 10.1002/sim.5869.

[33]. T. J. VanderWeele, "Mediation analysis: A practitioner's guide," Annual Review of Public Health, vol. 37, pp. 17–32, 2016, doi: 10.1146/annurev-publhealth-032315-021402.