KEYS TO STAYING SHARP

Interventions to Attenuate Cognitive Decline

Supported by:

The National Institute on Aging

R01 AG056428

Modifications:
(Any modification to the protocol after study investigators meet and approve it will be annotated. The annotation should note the exact words that are changed, the location in the protocol, the date the modification was approved by the investigators, and the date it became effective, if different from the approval date.)

Table of Contents

1 SYNOPSIS

1.1 Study Title

Interventions to Attentuate Cognitive Decline: Keys to Staying Sharp.

1.2 Objectives

The primary objectives are:

- To examine the efficacy of piano training to improve central auditory processing (CAP), cognition, and everyday function among older adults with and without mild cognitive impairment (MCI).

The secondary objectives are:

- To examine the moderating effects of MCI on piano training efficacy.
- To explore mediators of intervention effects.

1.3 Design

The design is a two arm randomized clinical trial examining the efficacy of piano training relative to music listening instruction to improve CAP, cognition, and everyday function in older adults with and without MCI across two time points (baseline and immediate post-test). This is a phase II trial to test the efficacy of promising behavioral intervention (i.e., piano training) in a research setting. The primary outcome of interest is cognition[1].

1.4 Outcomes

We will quantify the effects of piano training on CAP, cognition, and everyday functional performance.

CAP measures will include: Time Compressed Speech 65%, Words-in-Noise, Dichotic Digits Test, Dichotic Sentence Identification, and Adaptive Tests of Temporal Resolution.

Cognition measures will include: Verbal Fluency Test (phonemic fluency, category fluency, and category switching), Trail Making Test, and Digit Coding.

Everyday Function measures will include: Timed Instrumental Activities of Daily Living and Test of Everyday Attention.

1.5 Interventions and Duration

Piano Training. Piano training will consist of basic piano technique, dexterity exercises, piano literature, and music theory. At each bi-weekly training session, participants will be trained in groups and will be expected to

perform all technique (scales/finger dexterity exercises), piano repertoire (Alfred Basic All-in-One Method), and complete music theory assignments. Each class session will be structured with a short review followed by an intense focus on new skill development (e.g., scales; chord progressions) and concept formation (e.g., intervals). At least one five-minute break will occur in each session.

Music Listening Instruction Condition. Participants randomized to the music listening instruction condition will engage in music listening and appreciation. The participants read about, listen to, view diagrams of, and answer questions about the music. Each lesson will cover chapters from *Music Listening Today* over 90 minutes. Each lesson will be structured as follows: Overview of chapter material and group listening activity, participants read over chapter on their own, active listening activity, five question quiz, overview of next chapter material and group listening activity, participants read over chapter on their own, active listening activity, and five question quiz. At least one five-minute break will occur in each session. This structure is equivalent to the switching between ensemble/individual approach to teaching in the piano training condition.

The two training conditions will be equivalent in terms of frequency and duration of each session (90 min/day, two days/wk, 10 weeks) and social contact (led by trainer and conducted in groups of up to 10 persons). Both conditions will be described as "music training".

Sample Size and Population

Our goal is to have up to 360 participants complete the study. Individuals with normal cognition and those with a clinical diagnosis of MCI will be recruited for this study. Inclusion and exclusion criteria are detailed below.

2 STUDY OBJECTIVES

2.1 Primary Objective

The primary objective is:

- To examine the efficacy of piano training to improve CAP, cognition, and everyday function among older adults across baseline to immediate post-test. We will assess the effects of piano training (intervention) as compared to music listening instruction (active control) on CAP, cognition, and everyday function. The primary outcome of interest is cognition.

  Hypotheses: The piano training group will outperform the music listening instruction group on measures of CAP, cognition, and everyday function immediately post-intervention.

  Rationale**:** The efficacy of a novel intervention will be established relative to an active control group. There is compelling evidence that CAP deficits contribute to cognitive difficulties [2-5] and are a risk factor for dementia [3, 6, 7]. We propose that enhancing CAP may be particularly important to improve cognitive function. Formal music training is longitudinally associated with enhanced CAP and successful cognitive aging. Piano training is a feasible, potentially efficacious intervention to improve older adults' CAP and cognition.

2.2 Secondary Objectives

The secondary objectives are:

- To examine the moderating effects of MCI on piano training efficacy. Moderating effects of MCI status will determine if piano training is effective for older adults with and/or without MCI.

  Rationale: It is possible that piano training is effective prior to cognitive decline, but not in those with MCI. It is also possible that cognitive and/or functional abilities may be differentially enhanced among those with and without MCI. Delineating the effects of piano training among those with and without MCI will discern who is most likely to benefit from piano training, and for what specific outcomes. We do not have a directional hypothesis for this objective.

- To explore mediators of intervention effects.

  Hypothesis**:** Grounded in theory[8] and prior research[9-11], we hypothesize that piano training will improve CAP and cognitive performance, leading to functional improvements. We expect that enhanced CAP will mediate cognitive gains. We further hypothesize that cognitive gains will mediate functional improvements.

  Rationale**:** Crucially, we must discern the underlying mechanisms of effective interventions in order to maximize training gains.

3 BACKGROUND

Cognitive impairment and dementia are the most feared signs of aging[12-14]. In addition to the obvious health and quality-of-life ramifications for older adults, there are high economic costs (e.g., subsidizing nursing home care; lost productivity of family caregivers), when older adults can no longer live independently[15]. Thus, prevention of cognitive and subsequent functional decline among older adults is of critical importance to public health. Dementia is the most expensive medical condition and increases in prevalence with age[16, 17]. Given that older adults are the fastest growing segment of the population[18], the prevalence of dementia (and its precursor, mild cognitive impairment-MCI) has been increasing rapidly[19]. Therefore, finding effective strategies for intervention to prevent or delay dementia is imperative to improve public health. The contribution of the proposed research is expected to be the identification of a novel and effective cognitive intervention to counter age-related cognitive and functional decline, which could potentially delay dementia onset. We will elucidate the underlying mechanisms of effective interventions to facilitate maximizing cognitive gains and improving everyday function. The contribution of this research will be significant in that if an intervention could delay the onset of Alzheimer's disease by only one year, there would be about 9.2 million fewer cases of the disease in 2050[20], substantially lessening the burden. Results will inform subsequent research and clinical practice by facilitating the design and implementation of effective interventions to attenuate cognitive decline and thereby improve public health.

Mild Cognitive Impairment (MCI). Cognitive status has been conceptualized on a continuum from normal cognition to clinically ascertained dementia such as Alzheimer's disease. MCI represents a transitional phase between normal aging and dementia[21, 22]. MCI patients have higher risk of developing dementia[12, 23] with conversion to dementia of ~12% per year[23] (although not all individuals progress across this continuum). Among individuals with and without MCI, cognitive interventions may enhance cognition and maintain everyday functioning, thereby delaying dementia onset[24]. A paucity of research and methodological inadequacies have limited the implementation of effective cognitive interventions in MCI[25, 26]. One avenue to enhance cognition is the facilitation of central auditory processing (CAP) at initial perceptual stages[5], which may potentially be achieved by piano training. Our scientific premise is that piano training will enhance CAP, resulting in improved cognition and subsequently everyday function, thereby delaying the onset of dementia. The primary outcome of interest is cognition.

Central Auditory Processing (CAP). One early indicator of cognitive impairment is impaired auditory processing. Auditory processing and decoding of sound occurs at multiple levels of the central auditory nervous system. Central auditory processing dysfunction is characterized by deficits in using supra-threshold sound, such as poor monaural speech-in-noise perception, binaural speech processing, sound localization and lateralization, and auditory processing speed [27]. When CAP is impaired, information received by the cognitive systems is distorted and incomplete, causing information processing difficulties. Auditory processing thus requires greater cognitive resources, negatively affecting memory and executive function [8]. Decades of research have established that CAP is fundamentally linked with memory and executive functioning[4, 5, 28, 29], which are the primary deficits in MCI and dementia[30, 31]. In support of this premise, CAP longitudinally predicts MCI and dementia such as Alzheimer's disease[32-34]. Prior research (e.g., [35, 36]) and our pilot data confirm that measures of CAP are strong, independent predictors of cognitive function (e.g., memory, executive function), and that adults with MCI show significant CAP deficits[37, 38]. Deficits in CAP are multifaceted and represented by difficulties with, for example, auditory speed of processing, monaural speech-in-noise recognition, binaural speech processing, sound localization and lateralization, and auditory pattern recognition[5, 39]. Note that auditory speed of processing, a leading indicator of cognitive aging, is one of the key elements of CAP[40], as adequate sensory functioning is fundamental for normal speed of processing[41]. When CAP is impaired, information received by the cognitive systems is distorted and incomplete, causing information processing difficulties. Auditory pro-

cessing thus requires greater cognitive resources, negatively affecting memory and executive function[8]. Therefore, impaired CAP impedes cognitive performance[41-45]. In addition to behavioral data, neuroanatomical evidence documents that deficits in CAP[46] are related to anterior temporal lobe atrophy and reduced glucose metabolism[46-50], brain changes evident in MCI and Alzheimer's disease[51]. Neuroimaging further indicates a link between sensory function and cortical integrity, and that decreased volume in the auditory cortex cascades to affect higher-order cognitive processing[52-54]. In older adults, difficulties with CAP are associated with modified patterns of neural activation as well as with reduced gray matter volume in the auditory cortex. Wong et al. and others[55, 56] demonstrated that older adults show less activation in the auditory cortex and increased activation in the prefrontal cortex to perform cognitive tasks due to declines in CAP.

Although CAP is a strong longitudinal predictor of cognitive decline, MCI, and dementia[3, 34, 35, 57], prior research has not examined the effects of interventions targeting CAP on functional abilities of older adults with and without MCI. Re-establishing and/or maintaining adequate CAP may be the first crucial step in efforts to improve cognitive abilities of older adults. Given that multiple studies have shown music experience is strongly associated with better CAP[58-61], it is likely that piano training enhances CAP, leading to better cognition (See **Piano Training Rationale** below).

Theoretical Background. We hypothesize that improving brain function at early perceptual levels (i.e., CAP) may be optimal to attenuate cognitive and functional decline and potentially curb dementia prevalence. This premise is grounded in the information degradation[8] theory. The information degradation theory[8, 62] posits that age-related changes in the brain cause initial sensory/perceptual processing errors that lead to difficulties with downstream information processing (i.e., CAP) and cognition (i.e., memory/executive function). According to this and similar theories, neurophysiological decline leads to reduced brain efficiency such that the speed and/or quality of processing (e.g., fidelity of representations, speed of neural transmission) is impaired due to poor perceptual encoding. Thus, cognitive interventions should target improving initial perceptual processing (i.e., CAP) such that cognition can improve[63]. Accordingly, cognitive interventions that target basic perceptual processing may be most effective. Furthermore, cognitive gains will be mediated by improvements in perceptual processing (i.e., CAP). Our recent analyses support this assertion, indicating that intervention targeting perceptual processing is more effective to curb dementia prevalence than other approaches[64]. The proposed study will advance the field by examining CAP as a mechanism of piano training.

Piano Training Rationale. Engagement in learning challenging new skills, such as learning to play the piano with training, is a promising cognitive intervention approach that is likely more effective than cognitive stimulation[65]. Piano training can be defined as systematic sequential instruction to acquire piano skills that includes technique (scales/arpeggios/chords), music theory, and piano repertoire. Verghese et al. observed that older adults who reported playing musical instruments were 69% less likely to develop dementia[66]. Increasing neuroscience evidence indicates positive, longitudinal associations of music training with enhanced structures and function in the brain[67, 68]. Thus, music training (piano training, in particular) has been proposed as a viable means to attenuate age-related cognitive decline[69, 70]. Piano training may be an efficacious cognitive intervention because it requires the simultaneous coupling of sensory processing, motor skills, and complex cognitive operations to be executed with temporal precision[71]. We propose that piano training is a promising cognitive intervention because:

a- Piano training enhances CAP. Our central premise is that improving CAP may be the first crucial step to augment older adults' cognition. Numerous studies insdicate that music experience is strongly associated with better CAP (e.g., [58, 59-61]). CAP is likely enhanced by piano training due to the multicomponent demands on complex temporal, sequential, and sensorimotor processing. Correlational studies indicate superior CAP for adults with formal music training compared to non-musicians[72-74]. Further, Co-I Lister's studies show that adult musicians have enhanced CAP, indicated by faster neurophysiological processing of sound, than their nonmusician counterparts[9, 75]. Similarly, older musicians do not show age-related neural timing delays, indicat-

ing better CAP[76]. Zendel and colleagues also demonstrated that CAP (i.e., auditory processing speed and speech-in-noise perception) is less suspectible to age-related decline in musicians[59]. We propose to experimentally examine the effects of piano training on older adults' CAP in a RCT.

b- Piano training enhances brain structures and function. Epidemiological data show that music training is associated with enhanced structures and function of the brain. Scientists have hypothesized that piano training engages frontostriatal cerebellar circuits, including the motor cortex and basal ganglia, and will thus result in cognitive improvements among older adults[71, 77, 78]. Research demonstrates that long-term piano training is associated with greater activation in the cerebellum, an area associated with higher cognition (e.g., executive function)[79]. Pianists show higher white matter integrity and greater gray matter density in motor areas of the brain than non-musicians[79]. Neuroimaging research indicates specificity in structural brain differences subsequent to piano training[80]. Given that piano training is longitudinally associated with improved brain structures and function, it may attenuate age-related cognitive decline.

c- Piano training enhances cognition. Piano training is associated with improved cognitive performance[81, 82, 83-85]. Hanna-Pladdy and Gajewski demonstrated that as compared to non-musicians matched on age and education, older adult musicians performed better on memory (letter number sequencing, verbal learning test *d*s=0.47-0.53) and executive function measures (phonemic fluency *d*=0.59)[86]. Musical training was a stronger predictor of older adults' cognitive performance than lifestyle activities[87]. Seinfeld and colleagues found that older adults randomized to piano training showed improved speed of processing (*d*=0.71) and executive function (*d*s=1.70-1.85)[67]. Co-I Bugos published two pilot randomized trials indicating small to medium effects of piano training relative to controls for improved speed of processing and executive function. Improvements in speed of processing were sustained three months after training (for details, see **Preliminary Studies-*Piano Training*** below). Combined with our pilot study, these data demonstrate enhanced speed of processing and executive function among older adults subsequent to piano training relative to controls[81, 82].

d- Piano training enhances older adults' mood and quality of life. In our experience, older adults are eager to participate in piano training studies and find the experience to be enjoyable. Seinfeld and colleagues found that older adults completing piano training reported less depression, improved mood, and better quality of life[67]. Piano training may be a more enjoyable activity than other types of cognitive intervention. Further, piano training could be immediately implemented in the community if it is efficacious.

In summary, we chose piano training because it is a novel cognitive engagement approach that shows promise to enhance CAP and cognition. Although prior epidemiological research and our pilot studies indicate the potential efficacy of piano training[81, 82, 83-85], statistically powered experimental studies are lacking. Thus, we propose to examine the efficacy of piano training in a phase II randomized trial. Per Onken's model of behavioral interventions, we are examining efficacy in a research setting[1].

4 STUDY DESIGN

A randomized clinial trial (RCT) with two arms (piano training and music listening instruction) across two levels of cognitive functioning (cognitively intact older adults with and without MCI) will be rigorously conducted over three project years. Potentially eligible participants will complete telephone screening, in-person baseline, and if needed, a clinical assessment to diagnose MCI, will be randomized to complete piano- or music listening instruction, and will complete immediate post-training assessments. This is a phase II trial to test the efficacy of a promising behavioral intervention (i.e., piano training) in a research setting[1].

5 SELECTION AND ENROLLMENT OF PARTICIPANTS

5.1 Inclusion Criteria

Inclusion criteria will be:

- aged 60 years or older
- willingness to provide informed consent
- willingness to complete up to 23 study visits at USF including attending in-lab intervention two times a week for a three-month period with the goal of completing 20 sessions.
- ability to speak and understand English
- Montreal Cognitive Assessment score of 20 or higher (score 20 to 30 inclusive)
- intact vision (binocular near visual acuity of 20/50 or better tested with a standard near visual acuity chart)
- adequate hearing acuity (no greater than a moderate hearing loss [thresholds <70 dB HL] in the mid-frequency range [1000, 2000 Hz] in at least one ear as determined by a standard pure tone hearing evaluation)
- Music Reading Assessment score of 18 or lower (score 0-18 inclusive)
- ability to understand study procedures and comply with them for the length of the study in the tester's opinion (and other study personnel opinion who interact with participant, such as the study physician)

5.2 Exclusion Criteria

Exclusion criteria will be:

- moderate or worse depressive symptoms (GDS short form score >=5)
- previous participation in USF Cognitive Aging Lab or Music Research and Testing lab[1] intervention studies
- previous participation in 10 or more hours of a computerized cognitive intervention computer programs (e.g., Lumosity, Posit Science Brain Fitness, InSight, or Brain HQ; Lace, CogMed, CogniFit, Happy Neuron, Dakim, DriveSharp or Staying Sharp by AARP programs)
- currently enrolled in another research study
- planning on being away for two or more weeks during the next five months (recruit later)
- undergoing chemotherapy or radiation treatment or planning surgeries or other procedures requiring anesthesia within the next five months (recruit later)
- four or more years of formal music training such as private lessons or group lessons on a specific instrument
- ability to read music on two or more of the following clefs: Treble clef, Bass clef, Alto clef
- four or more years of playing any one musical instrument
- currently practicing or participating in any music activities- such as music performance or music reading or music lessons or music related courses
- difficulty and pain in moving hands or fingers, or neuropathy (e.g., numbness or tingling) affecting hands, which interferes with the ability to use a keyboard[2], or tremor in either hand, or missing any finger or portion of a finger
- self-reported diagnosis of dementia, stroke, serious brain injury or neurological disorder
- diagnosed by a physician or nurse with a TIA that occurred within the last 18 months
- inability or unwillingness to give written informed consent at baseline
- Clinical Dementia Rating Scale score of 1 or greater
- Clinical diagnosis of dementia or other disorder that in the study physician's opinion would limit the persons ability to participate in the study or benefit from the interventions

[1] Although the "Music Research and Testing Lab" was omitted from version 1, this criteria was included in telephone screening from the study onset. Potential participants were excluded if they reported prior participation in Music Research and Testing Lab studies.
[2] We added this clarification, "that interferes with ability to use a keyboard" to the screening process.

5.3 Enrollment Procedures

Participants will primarily be recruited through the community, the Cognitive and Neurophysiology of Aging Lab registry, and mailing list of older adults residing in the area. We may also use the Byrd Alzheimer's Institute Community-Based Memory Screening program and the Department of Psychiatry and Behavioral Neurosciences Memory Disorders Clinic (MDC) to enroll participants in the study.

Recruitment and enrollment will occur across a pilot and five or six replicates. Pilot replicate data will only be included in interim or final analyses if no significant changes to the study procedures or this protocol are made after the pilot replicate is completed. Any changes to the protocol will be noted with the effective date and justification.

Throughout the study period at least one Research Assistant in the Cognitive and Neurophysiology of Aging Lab will make recruitment phone calls and answer potential participant emails/calls. Individuals who have expressed interest in our research will be telephone screened (i.e., inclusion/exclusion questionnaire) by lab staff. The study script should be followed for describing the study and answering any participant questions. Screeners will emphasize to the participants that they are committing to complete up to 23 in-person study visits (i.e., 3 testing sessions and 20 intervention sessions) over the next five months. Once eligibility and interest has been assessed, the final outcome is coded. The final recruitment outcome will be coded as 1) Scheduled; 2) Not Scheduled and Will Not Try Again; or 3) Not Scheduled, Will Try Again.

If the participant is not eligible because they are planning on being away or unavailable for more than 2 weeks in the next 5 months, or are planning on having major surgery or undergoing any type of anesthesia, chemotherapy, or radiation treatment in the next 5 months, we will attempt to recruit them for the next replicate.

The study coordinator will oversee recruitment and serve as a primary contact for study participants.

Those deemed initially eligible from telephone screening will complete an in-person screening at the Cognitive and Neurophysiology of Aging Lab. This in-person screening will include the Montreal Cognitive Assessment (MoCA), Geriatric Depression Scale, vision and hearing screenings as well as the Music Reading Assessment (MRA) to further confirm eligibility (See Inclusion/Exclusion above for details).

Those potentially eligible based on completion of these measures will further complete baseline cognitive, CAP, and everyday functional measures.

Those eligible who score 26[3] or better on the MoCA will be enrolled in the study and referred to the study coordinator for randomization.

Those potentially eligible who score between 20 to 25 on the MoCA who were not referred by the MDC or study physician will also complete a clinical evaluation visit. A cut-point of 20-25 on the MoCA is indicative of probable MCI, while less than 20 is suggestive of dementia[88].

[3] The MoCA cutpoint was changed from 27 to 26 to use the more traditional and accepted value for impairment. Those who score between 20 to 25 will complete the clinical evaluation, rather than those who score 20 to 26.

The purpose of the clinical evaluation is to determine if participants have MCI or dementia. The clinical evaluation visit will include subtests of the National Alzheimer's Coordinating Center Uniform Data Set (NACC) neuropsychological battery and an exam by a study physician. Subtests of the NACC battery will be used to determine MCI status and exclude dementia. The NACC battery subtests to be administered include the Clinical Dementia Rating Scale, the Multilingual Naming Test, the Benson Complex Figure Copy immediate and delayed subtests, the Craft Story 21 immediate and delayed recall subtests, Functional Assessment Scales, and Clinical Diagnosis Form, completed by a study physician who is a geriatric psychiatrist.

An exception to the above procedure will be for patients with MCI who were referred from the USF Health Memory Disorders Clinic (MDC) or a study physician who have already undergone clinical evaluation in the past 6 months. Eligible and enrolled participants referred from the clinic with prior MCI diagnosis and evaluation in the clinic completed within the past 6 months will be referred to study coordinator for randomization.

In replicate 1 and later replicates, we plan to request lab results from the past year from those who score 20-25 on the MoCA. The purpose is to rule out medical causes of cognitive decline that are treatable. If copies of lab results are not obtained at the baseline visit, we will request that the participant complete a release of medical records. If we do not receive medical records by the clinical evaluation, and the study physician recommends it, we may draw blood and order lab tests that may include:

- CBC
- Metabolic panel
- Thyroid Stimulating Hormone (TSH)
- B12
- D
- Folic Acid

Any results not within normal limits will be reviewed by a study physician and communicated to the participant.

Those without any clinically significant abnormalities that would likely interfere with their ability to benefit from intervention[4] will continue in the study.

5.4 Describe consent procedures.

Written informed consent from participants will be obtained prior to in-person screening at the first study visit. Research Assistants who have completed human subjects training and are approved by the Institutional Review Board will explain the study procedures, offer to read the consent form to the participant (and will do so, if desired), answer any questions, and obtain informed consent. A flow chart that details the study visits and timeline will be provided to the participant to help explain the study procedures. Participants will be encouraged to take time to consider whether they wish to participate in the study. The person obtaining consent will leave the room to provide the participant time to read and consider the informed consent statement. The participant will be provided an opportunity to ask questions prior to signing the consent (as well as throughout the study). A copy of the consent form will be provided to the participant.

[4] We added "and communicated to the participant", and we added the clarification, "that would likely interfere with their ability to benefit from intervention".

All signed consent forms will be kept in a locked filing cabinet separate from the data to ensure confidentiality. In the data quality control process, each consent form will be reviewed to ensure accurate completion and protocol adherence.

- Informed consent will be an agreement between the investigators and each participant having the capacity to understand and make an informed decision. Consent will be obtained prior to each potential participant's participation in this study.

- Each individual participating in this study will be made aware of the fact that his or her participation involves research and the intent of the research, the expected duration of their participation and a description of the procedures that will be followed.

- Each participant will be made aware of the reasonably expected benefits he or she might receive, as well as any risks or potential discomfort that are involved.

- Each participant will be made aware of alternative procedures that are available to him or her.

- Each participant will be made aware that his or her records will remain confidential, but that the NIH, IRB, and OHRP have the right to inspect his or her records.

- Each participant will be told that his or her participation in the clinical study is voluntary, without force or influence from the investigators or research staff.

- Each participant will be given the name and method of contacting the appropriate person(s) to answer his or her questions about the research and in the unlikely event of research-related injury.

5.5 Describe the procedure for obtaining intervention group assignment

The study coordinator with the statistician will randomize participants stratified by MCI status. Participants with and without MCI will be randomized in a 1:1 ratio to either piano training or the active control condition of music listening instruction. To ensure balance throughout enrollment, randomization will be stratified by MCI status (present vs. absent) and conducted in varying blocks of participants as designed by Co-I Ji. Dr. Ji will generate the stratified permuted random blocks using R. Dr. Ji will maintain the random seed and the computer codes and will send the randomization list to the clinical trial coordinator. The treatment allocation will be denoted as A or B and recorded on the randomization form by the clinical trial coordinator. The study coordinator will control access to the randomization list and know whether A (B) stands for intervention (control). Study investigators Edwards, Lister, and Andel will remain blind to participant condition. The testers will remain blind to participant condition. The trainers (including Co-I Bugos) and clinical trial coordinator will be aware of participants' randomized condition.

5.6 STUDY INTERVENTIONS

5.7 Interventions, Administration, and Duration

There are two arms in this study including group piano training (intervention) and music listening instruction (active control). Both of these arms are described to participants as "music training". The detailed procedural protocol for both training arms can be found here: P:\CBCS\CSD CAL\Study 15 Music Training NIA\Trainer Docs\Study 15 Music Training Procedure Protocol…. or upon request to the PI.

The content protocol for the **piano training** arm can be found here P:\CBCS\CSD CAL\Study 15 Music Training NIA\Trainer Docs\Piano training Manual_JB_Copyright_2018.docx or upon request to Co-I Bugos.

The content protocol for the **music listening instruction** arm can be found here P:\CBCS\CSD CAL\Study 15 Music Training NIA\Trainer Docs\Music Listening PowerPoints.zip or upon request to the PI.

Participants will be scheduled by the trainers or study coordinator for in-lab intervention phase sessions in groups of up to 10 individuals by randomized condition. We are limited to 10 persons per class due to the physical space available for the music listening condition. The piano and music listening labs are located on the USF campus and both can accommodate 10 participants in a group intervention session. Scheduled participants will be sent an appointment letter with directions and a map (by mail if visit is 8 or more days, by email if visit is within the next 7 days). Participants will be reminded one to three days prior to the first intervention visit by telephone by a Research Assistant.

Both conditions will attend training two times a week for 10 weeks with the goal of completing a total of 20 sessions before post-testing. Participants will be encouraged to complete training two times a week without missing two or more consecutive sessions. Make up sessions will be held for each condition weekly, as needed[5]. The training phase can be extended up to a maximum of 15 weeks, if needed to complete 20 sessions. Participants will be encouraged to complete post-test regardless of intervention adherence. We will code participants as completing training if 16 or more sessions are completed (16 to 20 sessions inclusive). All training sessions must be completed within 15 weeks of training start date. The Trainers will confirm or reschedule the participant to complete an immediate post-intervention assessment within 2-60 days of their final in-lab intervention session (i.e., training completion date), or within 18 weeks of baseline if none of the the intervention phase was completed. See further details below in **Intervention Phase** and **Retention.**

5.7.1 Prohibited Interventions

Participation in cognitive interventions aimed toward enhancing or maintaining cognitive abilities will not be allowed. This includes computerized cognitive training or stimulation exercises aimed at countering cognitive decline. Participants will be instructed not to participate in such activities during the entire study period. Formal music training, such as private music lessons or group music classes or courses related to music (as example voice instruction or musical instrument instruction or a music appreciation class), outside of the assigned activities will not be allowed. Individuals who participate in such interventions will be dropped from the study.

[5] The following sentences were omitted: Each of the conditions will have a textbook and will be assigned weekly homework. Participants will be asked to report the amount of time they spend outside of class completing such activities.

Only data points prior to such participation will be included in analyses. Participant will be deemed ineligible at the time such an incident occurs.

If the participant opts to undergo a surgical procedure that requires general anesthesia or chemotherapy or radiation, we will attempt to complete post-test prior to the surgery or procedure. If the participant undergoes general anesthesia, chemotherapy, radiation, experiences a head injury, stroke, or heart attack, the participant will be withdrawn from the study. All available data points prior to the incident will be used in analyses. The participant will be deemed ineligible at the time of the incident.

Participants will be provided a post-test appointment letter that will include instructions regarding these items when scheduled for post-test. Also, the trainers will remind participants of these instructions on the first day of training.

5.8 Adherence Assessment

To address the study aims, we will use intent-to-treat analyses as detailed in section 9.5 below. We will report the percent of participants who were adherent to the assigned piano training or music instruction exercises. Participants will be considered adherent if they complete at least 80% of the assigned sessions (at least 16 of 20 sessions). The Research Assistant Trainers will record dates that participants attended in-lab training. Participants will be allowed up to 15 weeks to complete the assigned exercises between baseline and post-test.

Primary data analyses will be intent-to-treat. Sensitivity analyses will further examine effects of piano training relative to music instruction among those who were adherent (i.e., who completed 16 or more hours) in both conditions.

6 CLINICAL AND LABORATORY EVALUATIONS

Schedule of Evaluations

Evaluation	Measure type	Telephone Screening (≤ 60 days)	Baseline	Clinical Evaluation (≤ 45 days)	Randomization ≤3 wks	Intervention Up to 15 wks	Post-Test >2 days and ≤=60 days
Informed Consent	R		X				
Inclusion/Exclusion Questionnaire	I/E	X	X	X			X
Geriatric Depression Scale (GDS)	I/E, C		X				
Demographic Data	C		X				
Montreal Cognitive Assessment (MoCA)	I/E, C		X				
Memory Screening History	IV		X				
Hearing and vision screening Pure-tone Hearing Thresholds Speech Recognition Threshold Near Visual Acuity	I/E, C		X				X
Music Reading Assessment	I/E		X				
Central Auditory Processing Time Compressed Speech 65% Words-in-Noise Test Dichotic Digits Test Dichotic Sentence Identification test ATTR	O		X				X

Evaluation	Measure type	Telephone Screening (≤ 60 days)	Baseline	Clinical Evaluation (≤ 45 days)	Randomization ≤3 wks	Intervention Up to 15 wks	Post-Test >2 days and ≤=60 days
Cognitive Assessment Battery Trail Making Test Digit Coding Verbal Fluency - BHR, Clothes/ Girls' Names, Vegetables/ Musical Instruments (baseline) - FAS, Animals/ Boys' Names, Fruits/Furniture (post-test)	O		X				X
Everyday Functional Assessment Timed IADL Test Test of Everyday Attention	O		X				X
Questionnaires Health Questionnaire Medication Audit Health Changes	SC		 X X				 X
Adverse Events	R		X	X		X	X
Clinical Evaluation NACC assessments: -Clinical Dementia Rating Scale -Craft Story 21 Recall - Immediate and Delayed -Benson Complex Figure Copy and recall -Multilingual Naming Test -Functional Assessment Scale -Clinician Diagnosis form	SC			X			

Evaluation	Measure type	Telephone Screening (≤ 60 days)	Baseline	Clinical Evaluation (≤ 45 days)	Randomization ≤3 wks	Intervention Up to 15 wks	Post-Test >2 days and ≤=60 days
Lab Assessments (upon IRB approval) CBC Metabolic panel Thyroid Stimulating Hormone (TSH) B12 D Folic Acid MRI (if funding acquired)	IV			X			
Randomization	IV				X		
Advanced Measures of Music Audiation (AMMA)- Day 1 of Training	E					X	
Basic Piano Measure (BPM) OR Music Listening Measure (MLM)- At days 1 and 20 of training	E					X	
General Self-Efficacy and Music Performance Self-Efficacy (At days 1 and 20 of training)	E					X	
Intervention sessions	C					X	
NICT Expectations Questionnaire (Between days 17 and 20 of training)	IV, C					X	
Tester blinding assessment	IV						X
C=covariate, E=exploratory, I/E=inclusion/exclusion, IV=internal validity, O=outcome, R=regulatory, SC=sample characterization							

6.1 Timing of Evaluations

Prior to Randomization: The baseline visit should be scheduled within 60 days of the telephone screening interview. Informed consent must be obtained first at the baseline study visit prior to any data collection. The GDS, MoCA, Music Reading Assessment, as well as hearing and vision screenings must be completed at baseline, prior to the clinical evaluation. The central auditory processing measures, cognitive assessment battery, everyday functional assessments and health questionnaire should be completed at baseline, but missing items can be administered at the beginning of the clinical evaluation, if needed. The clinical evaluation should be completed within 45 days of the baseline visit completion. The NACC assessments will be completed at the clinical evaluation. All the aforementioned measures for the baseline and clinical evaluations should be completed prior to randomization and must be completed prior to the training content being administered. It will be allowable to administer any of the aforementioned measures that are missing from the baseline or clinical visits after randomization, but administration will not be allowed once the participant has received any piano or music listening training.

Prior to Training Content: Any biomarker assessments must be completed prior to training content (blooddraw, mri, etc.). The general and music self efficacy measures can be administered at either baseline or after randomization on day 1 of training, but must be administered prior to participants receiving any piano- or music listening- content training. If the basic piano or music listening measures are administered prior to randomization, the participant will be withdrawn from the trial and a protocol deviation will be noted. The pre-training basic piano and music listening measures must be administered after randomization and must be administered before training content is delivered. If any of the aforementioned baseline, clinical, or pre-training measures are administered after the training content began, a protocol deviation will be noted. Such measures will be coded as missing and will not be used in analyses.

Intervention Content: The AMMA should be administered on day 1 of training, but it will be allowable to administer the AMMA within the first three training sessions. The NICT questionnaire should be administered between days 17 and 20 of training, but can be administered at the beginning of the post-test visit if it not completed during the intervention phase. The NICT questionnaire must be administered prior to the post-test assessments of CAP, cognition, or everyday function. If the NICT questionnaire is administered after the post-test assessments of CAP, cognition, or everyday function a protocol deviation will be noted. In this instance the NICT questionnaire will be coded as missing and will not be used in analyses. The post-training basic piano measure and music listening measure must be completed after all training content is delivered and prior to the post-test assessment visit. The trainer should confirm with the study coordinator that all scheduled intervention sessions have been completed and the basic piano or music listening assessments are completed prior to the post-test visit. The post-test visit can be rescheduled, if needed, but must occur within 18 weeks of baseline if no intervention was completed, or within 60 days of training completion date (>= 16 sessions completed). Participant post-tests will be conducted after the intervention phase for those who complete between 1-15 sessions. [6]

[6] The sentence "Participant post-tests will be conducted after the intervention phase for those who complete between 1-15 sessions" and "(>= 16 sessions completed)" was added for clarification.

Post-Test Content: The post-training general and music self-efficacy measures must be completed after all training content is delivered, but may be completed at the beginning of the post-test visit, if needed. The post-test assessments of CAP, cognition, and everyday functional measures cannot be completed until after a delay of at least 24 hours after the final intervention session (i.e., training completion date). The post-test assessments of CAP, cognition, and everyday function should ideally take place within 2-30 days of the training completion date (>= 16 sessions completed). The maximum window for post-test is that it must be completed within 60 days of training completion, or within 18 weeks of baseline if no intervention phase was completed. Participant post-tests will be conducted after the intervention phase for those who complete between 1-15 sessions.[6]

6.1.1 Pre-Randomization Evaluations

Telephone screening is an initial interview conducted by Cognitive and Neurophysiology of Aging Lab staff to ascertain potential eligibility of the participants per the inclusion and exclusion criteria. Eligibility will be evaluated by the telephone screening. If ineligible the reason will be documented. Procedures for eligibility and enrollment are detailed above. Those potentially eligible will be scheduled to complete a baseline visit within 30 days. If the telephone screening questionnaire has not been completed within 60 days of scheduling baseline, it should be repeated.

A subject ID number will be assigned on the participants visit control sheet (VCS) by the tester at their first study visit. Informed consent will be obtained at the beginning of the first-in person visit, the baseline visit. Participants will be provided a flow chart that explains all of the study visits, duration of the visits, and the timing of the visits during the consent process. At the baseline visit, participants will complete the MoCA, GDS, hearing, and vision screenings, and MRA to determine eligibility. Those not eligible due to MoCA, GDS, hearing or vision scores will be referred to the appropriate professionals for treatment. Additionally, participants for whom outer or middle ear pathology is indicated through otoscopic examination will be referred to an otolaryngologist or other professional and will not continue in the study. Those participants who are eligible will be enrolled in the study and will further complete thorough assessments of CAP, cognition, and everyday function. A questionnaire assessing health will be completed. Research Assistant Testers who will remain blind to participant condition will complete these assessments. Participants will be referred to the study coordinator for randomization.

If the participant meets eligibility criteria, was not referred from the MDC or a study physician, but scores between 20-25 on the MoCA, a clinical evaluation will be performed by a study physician to ascertain diagnosis of mild cognitive impairment and ensure that clinical inclusion criteria are met and no exclusion criteria are evident. Reasons for ineligibility will be documented. This clinical evaluation should be completed as soon as possible following the baseline visit and should take place within 45 days after the baseline visit.

6.1.2 Intervention Phase

Within three weeks of baseline or clinical evaluation phase completion for each replicate, the Clinical Research Coordinator and/or Research Assistant Trainers will randomize enrolled participants in collaboration with Dr. Ji and schedule their first intervention visit. Participants will complete randomized exercises in groups of up to 10 individuals at in-lab sessions that are 90 minutes in duration, occurring two times a week over the subsequent ~3 months. The goal will be for 95% of the training sessions to be conducted in groups of 3-10 persons. Clas-

ses will be arranged such that those with MCI primarily train together within randomized condition. The intervention phase will continue until 20 total sessions are completed or 15 weeks have passed. The trainers will confirm or reschedule the participant to complete an immediate post-intervention assessment. The trainers will make certain that all scheduled intervention sessions are completed prior to the post-test visit and that no training sessions are scheduled after the post-test visit. Fidelity will be assured as described below (See **Quality Assurance** below**).**

At the first training sessions the trainers will administer the Advanced Measure of Music Audiation (AMMA), General Self-Efficacy Form, and the Music Performance Self-Efficacy form. Those in the music listening control group will complete the music listening measure while those in piano training will complete the basic piano measure. The training protocols detail all of the procedures and content protocols detail the training arms. Please see: P:\CBCS\CSD CAL\Study 15 Music Training NIA\Trainer Docs\Study 15 Music Training procedure protocol.

Between sessions 17 and 20, the trainers will administer the The Cognitive Training Expectations NICT questionnaire to examine participants' attitudes and expectations about the potential effects of the intervention completed. If a participant does not complete this questionnaire during the intervention phase, it may be completed at the beginning of the post-test visit.

At the participants' last training visit, the self efficacy measures will again be completed, and the trainer will verify that participants learned basic musical concepts and skills introduced in the program with at least 70% accuracy. If participants score less than 70% correct on the Basic Piano Measure (for piano training condition) or Music Listening Measure (for music listening condition), as applicable to randomized condition, at their last scheduled training session and if they have completed less than 20 sessions, then the trainer will an attempt to schedule additional make-up training sessions to complete up to 20 sessions. The trainers will make certain that the post-test visit is rescheduled, if needed. All training sessions must be completed within 15 weeks of training start date.

6.1.3 Post-Test

At the post-test visits, cognition-,CAP- and everyday function outcome measures will be re-administered by blinded Research Assistant Testers. Participants will complete the CAP, cognition, and everyday function measures administered at baseline. Alternative forms will be used, as available. The Tester will complete a form indicating whether or not they know the randomized condition of the participant to assess success of blinding and will confirm continued eligibility. A questionnaire will be administered to check for adverse events during the course of the study. Participants will be encouraged to complete post-test regardless of adherence. We will attempt to complete post test for up to 60 days after the training completion date (>=16 sessions completed) or 18 weeks post baseline if intervention was not attempted or completed. Participant post-tests will be conducted after the intervention phase for those who complete between 1-15 sessions.[6]

6.2 Overall Project Timeline

A detailed and up-to-date field schedule will be maintained to indicate the dates of each phase of the study for each of the pilot and five to six replicates. This schedule can be located at P:\CBCS\CSD CAL\Study 15 Music

Training NIA. Study data will be collected across a pilot and five to six replicates. The field schedule indicates the timeline for each phase of data collection during each replicate. Each replicate will last between 180 to 210 days. There should be at least two weeks between completion of the baseline/clinical phase and beginning of training phase to allow for randomization, scheduling, and coordination of intervention phase. The training phase will be 10-15 weeks in duration. We will attempt to complete post test for up to 60 days after training completion date (or 18 weeks post baseline if intervention was not completed).

6.3 Intervention Discontinuation Evaluations Retention

A CONSORT chart will be completed at the end of the trial. Participant status (e.g., screened, enrolled, refused, ineligible) will be recorded at each assessment. Reasons for ineligibility will be documented.

We will employ an intent-to-treat design as detailed in section 9.5 below. Participants who discontinue intervention will be encouraged to complete study visits and we will make every effort to follow and evaluate all enrolled participants.

6.3.1 Retention

We will employ several methods to encourage participation. In the consent process, we will stress that enrolling in the study involves a commitment to complete up to 23 in person visits: 3 testing visits and up to 20, 90-minute, in-lab training visits.

We find that asking participants to schedule their next appointment visit at the end of each study visit encourages continued participation. At each testing visit, we will provide a copy of the study flow chart with appointment times completed to remind them of the study visits and their progress.

Participants will be reminded the day before all scheduled testing visits by phone of their appointment time. They will be reminded 1-3 days prior to their first intervention (i.e., training) session.

We will indicate for each participant the best manner (e.g., email, phone, cell) and times to contact them to facilitate communication. We will also ascertain information for a secondary contact person for each study participant to facilitate follow-up. Such secondary contact individuals will be approached if we lose contact with the enrolled study participant with no response to phone calls or a letter after repeated attempts over a 90-day period.

At the first in-lab intervention visit, participants will complete a statement indicating their commitment to complete training sessions two times a week in order to complete 20 sessions. The statement will be signed by the participant and a witness. For any planned absences, an alternative make-up session will be scheduled. Typically in-lab intervention sessions are offered at regular times on Tuesday, Thursday or Monday, Wednesday, with make-up sessions scheduled on Friday. However, every effort will be made to accomodate participants' schedules (e.g., offering Saturday training sessions). The trainers will provide a schedule of in-lab training at the first intervention study visit. In week 8 of training, the trainers will confirm post-test visits for each participant. At this time, participants will again each be provided with a flow chart of the study visits with their spe-

cific study visit dates indicated. For any participant who could not be confirmed or rescheduled, the trainer will refer this person to the study coordinator who will assist.

Participants who miss an in-lab training session unexpectedly (i.e., without having a make-up session scheduled) will be contacted by the Research Assistant Trainer or lab staff within 3 days of the missed session. Participants will be requested to schedule a make-up in-lab training session as soon as possible. Participants will be reminded that their goal is to complete 20 sessions before post-test I visit. We will further encourage the participant to return for their post-test assessment visit regardless of training adherence. If we do not successfully make contact with the participant within the first three days, we will continue to attempt to contact them 2-3 times per week for three weeks. If we are still unable to reach the participants, we will send a letter requesting a reply. Secondary contact individuals will be approached if we lose contact with the enrolled study participant with no response to phone calls or a letter after repeated attempts over a 90-day period.

If participants are finding it difficult to attend the in-lab sessions, we will allow up to three sessions of training to be completed each week (1 session per 24 hours) until the intervention phase has ended or 15 weeks have passed since training began.

6.3.2 Documentation of MCI or dementia

For participants that score between 20-25 on the MoCA and who have not been referred by a study physician, the clinical evaluation visit will be completed. The clinical evaluation visit will be completed after baseline and before randomization. It should take place within 45 days of baseline, if possible. Participants will be requested to bring an informant (a relative or spouse or close friend who has regular contact with them who could report on their everyday activities). The informant will provide assent to participate in the study. We will allow informants to be interviewed by telephone, if unable to attend in-person. An evaluation will be performed by a qualified clinician to determine/confirm a diagnosis of MCI or dementia. After the participants complete NACC assessments, study physicians will evaluate the participant at the end of the clinical evaluation visit. The study physician will complete the Clinical Diagnosis Form and determine eligibility.

6.3.3 Concomitant Treatments

Participants will be allowed to undergo their usual treatment for health conditions.

6.3.4 Study Intervention Modifications

If at interim analyses, an effect size of d= 0.25 or greater reflecting improvements from piano training relative to controls is not observed on at least one of the cognitive outcomes measures (verbal fluency, Trails, Digit Coding) relative to controls, we will randomize the remaining participants to 25 sessions (i.e., 37.5 hours) instead of 20 sessions (i.e., 30 hours) of training.

6.3.5 Measures

- Eligibility will be assessed in-person by measuring near visual acuity, hearing, cognitive status (MoCA), Music Reading Assessment (MRA), and depressive symptoms (GDS). These variables and demographic information may also be considered as covariates in analyses.

- Near Visual Acuity will be measured at 40 cm using standard procedures with a Sloan letter chart, and the log minimum visual angle resolvable will be quantified [89].
- Hearing thresholds will be assessed using standard procedures (American Speech-Language-Hearing Association. (2005). *Guidelines for manual pure-tone threshold audiometry* [Guidelines]. Available from www.asha.org/policy.) and calibrated equipment from the USF Hearing Clinic. Otoscopic exam will be performed prior to inserting ear phones for pure tone testing. Pure tone air conduction thresholds in dB will be measured at 250, 500, 1000, 2000, 4000, 6000, and 8000 Hz in each ear. At post-test, these thresholds will be re-measured **only if** the participant reports a noticeable change in hearing since baseline testing (per Health Changes Checklist). Prescription for, use of (e.g., hours per day), and type of hearing aids will be recorded for each ear.
- Speech Recognition Threshold (SRT) will be assessed using standard procedures (American Speech-Language-Hearing Association. (1988). *Determining threshold level for speech* [Guidelines]. Available from www.asha.org/policy.) and calibrated equipment from the USF Hearing Clinic. An SRT will be measured for each ear. At post-test, these thresholds will be verified and presentation levels of all auditory tests will be adjusted if the SRT in either ear varies by more than +/- 5 dB.
- The Montreal Cognitive Assessment (MoCA) will be used to examine participants' cognitive status [88]. This assessment has good internal consistency, Cronbach's α=.83, and test-retest reliability, r=.92 [88].
- The Geriatric Depression Scale (GDS) will be used to screen depressive symptoms. The GDS has 92% sensitivity and 89% specificity to detect depression and has good test-retest reliability of r=.75-.81 [90, 91].
- Music Reading Assessment (MRA) [92] consists of a 40-item measure of basic music reading ability with 20 treble clef items and 20 bass clef items.

- Outcomes include binaural central auditory processing (CAP) measures of: Time Compressed Speech 65%, Words-in-Noise, Dichotic Digits Test, Dichotic Sentence Identification, and Adaptive Tests of Temporal Resolution. If any participant expresses discomfort with the presentation level, it will be adjusted to a comfortable level. The presentation level is recorded for each measure.

  - **Time Compressed Speech Test (TCS).** The TCS test assesses both auditory temporal processing and degraded speech understanding. It is a word recognition task in which the timing feature of the speech is digitally manipulated to resemble rapid speech [93, 94]. TCS performance worsens with age [95]. Auditec Recordings Northwestern University Auditory Test Number 6 recorded sentences (e.g., "Say the word jug") are available at both 45% and 65% compression (65% is faster). The listener must repeat the last word (test item), after the introductory sentence "Say the word", which preceeds every test item. The talker is female and approximately 4 sec of silence separate the stimuli. Participants are asked to repeat 50 words delivered through insert earphones at a level of 50 dB SL (re: SRT), and percent correct is calculated. Normative data on time compressed speech performance in adults have been described elsewhere [96]. For the present study, only the 65% condition will be used.
  - **Words-in-Noise Test (WIN).** The WIN is a test of degraded speech understanding or speech understanding in noise. It was developed as an instrument that quantifies the ability of listeners to understand speech in background multi-talker babble [97]. The stimuli are the Northwestern University Auditory Test Number 6 monosyllabic words with a static preceding sentence (e.g., "Say the word road") and a background of multi-talker babble. The talker is female and approx-

imately 3 seconds separate each sentence. Two 35-word lists are used; 5 words are presented at each of 7 signal-to-noise ratios (SNRs). The SNRs are presented in a decending manner, from 24 to 0 in 4-dB increments. The babble is maintained at a fixed level. Listeners must repeat the last word of the sentence, ignoring the babble. The WIN uses a modified method of constants to establish the SNR at which 50% correct performance is achieved on the materials. The 50% point is computed with the Spearman-Kärber equation [98, 99] [50% = 26 – (#correct)(0.8); "0.8" is the attenuation step size (4 dB) divided by the number of words per step (5)]. The WIN has been established as a reliable measure (intra-class correlation coefficient = 0.88) [100, 101].

- **Dichotic Digits Test (DDT).** The DDT is a test of binaural processing that uses a closed set of stimuli (numbers), thus having a relatively low linguistic load. The DDT consists of numbers between one and nine, excluding the two-syllable number seven (r=.79-.97)[102]. Two numbers are presented to the right ear while two numbers are presented simultaneously to the left ear, for a total of four numbers. The talker is male. Approximately four seconds of silence separate each set of four numbers, and approximately one second separates the individual pairs within a set of four. The participants are instructed to repeat all four numbers that were presented to them, in any order. Twenty-five sets of four numbers each are presented, for a total of 100 stimuli. The test is scored as a percent correct out of the 100 numbers presented.
- **Dichotic Sentence Identification (DSI).** The DSI test is used to assess binaural speech understanding in competition and also falls into the broad category of degraded speech understanding. The DSI stimuli are identical to those of the Synthetic Speech Identification Test (SSI) [103], recorded grammatically correct yet meaningless sentences (e.g., "Small boat with picture has become"). Unlike the SSI, for the DSI a competing sentence is delivered to the contralateral ear. The DSI is presented at a level of 50 dB SL (re: SRT) via insert earphones; the talker is male and approximately 8 sec of silence separate the sentences. Participants are required to select both sentences heard from a closed set list of 6 sentences, and results are scored as percent correct. Test-retest reliability of the DSI is high in older adults, r = .79 - .97. Deficits in DSI performance may be predictive of cognitive decline [104].
- **Adaptive Tests of Temporal Resolution (ATTR).** The ATTR[105] is a copyrighted and freely downloadable test of auditory temporal processing. Details regarding the ATTR have been published previously[105-107]. Briefly, the ATTR is used to measure gap detection thresholds using a three-interval, two-alternative forced-choice adaptive procedure targeting 70.7% correct gap detection. The stimuli used to define the silent gaps are ¼ octave narrow bands of noise (NBN) centered on either 1 kHz or 2 kHz. Before the silent gap, the NBN is 300 ms in duration. After the silent gap, the NBN varies randomly in duration between 250 and 350 ms. The participant is presented with three intervals of sound, a reference interval, a standard interval identical to the reference interval, and a target interval. In the reference and standard intervals, two NBNs separated by a 1-ms gap are presented. In the target interval, two NBNs separated by a gap of adaptively varying duration are presented. The reference interval is presented first and the standard and target intervals are presented in random order, following the reference interval. The participant's goal is to select the target interval from among the standard and target intervals. The participant is not allowed to select the reference interval. Two subtests of the ATTR were used in the present study, the within-channel subtest for which NBNs before and after the gap are both centered on 2 kHz and the across-channel subtest for which the NBN before the gap is centered on 2 kHz and the NBN after the gap is centered on 1 kHz. The ATTR was presented at a high, comfortable level, self-selected by each participant as described by Cox[108]. Reliability for the ATTR has been established with intraclass r =.58-.87[105].

- Outcomes also include cognitive measures: Verbal Fluency Test (phonemic fluency, category fluency, and category switching), Trail Making Test, and Digit Coding
    - **Verbal Fluency**[109]. Executive function will be assessed with phonemic, category, and category switching verbal fluency tasks. Phonemic fluency consists of three, one-minute trials during which the examinee names as many words as can be thought of that begin with a given letter, excluding proper nouns, numbers, and the same word with a different suffix will also be assessed (test-retest reliability=.88)[110]. Category fluency is designed to measure the speed and flexibility of verbal thought processes (r=0.82; sensitivity=.88; specificity=.96)[109, 111]. A switching condition is included that requires participants to alternate between sets (e.g. naming fruits and furniture) (r=0.51). Different versions of this measure will be administered at baseline and post-test. The BHR, Clothes/Girls' Names, and Vegetables/Musical Instruments will be administered at baseline while the FAS, Animals/Boys' Names, and Fruits/Furniture version will be administered at post-test. The recorded score is the number of correct items for each task.
    - **Trail Making Test Part A.** Speed of processing is evaluated using the Trail Making Test Part A. This test (r=.53-.64), requires participants to draw a line connecting a series of numbers in sequential order (1-2-3, etc)[112]. The recorded score is the time required to complete the task[113].
    - **Trail Making Test Part B.** Executive function is evaluated using the Trail Making Test Part B, which is a reliable measure (r=.54-.62) that requires participants to draw a line connecting a series of number and letters in alternating, sequential order (1-A-2-B, etc) (Strauss et al., 2006). The recorded score is the time required to complete the task[113, 114].
    - **Digit Symbol Coding.** Cognitive speed of processing is evaluated using Digit Symbol Coding [109]. This test consists of 135 blank squares that are paired with a number from one to nine. A reference key links each number to a different geometric figure. Using the reference key, participants have 120 seconds to copy the geometric figure assigned to the number into the blank boxes. Scaled scores are obtained by subtracting the number of errors from the number of correct responses. Lower scores indicate better performance. Psychometric properties are reported in the manual.

- Outcomes of everyday function include the Timed IADL and TEA tests
    - **Timed IADL** involves timed performance of five tasks encountered in daily life. The previously validated tasks[115] utilize real-world stimuli and represent five IADL domains, including communication (finding a telephone number in a phone book), finance (making change), cooking (reading the first three ingredients on a can of food), shopping (finding two items on a shelf of packaged foods) and medication management (reading the directions on a medicine bottle label). Scores are generated by combining the completion time and error code for each task per standard procedure. Task scores are combined into a single composite by taking the average of z scores computed for each of the five tasks after error correction. Test-retest reliability of the Timed IADL is r = .85 [115].
    - **The Test of Everyday Attention (TEA)**[116, 117] consists of eight sub-tests designed to measure different types of attention. The subtests are designed to mimic everyday tasks that encompass both visual and auditory domains. The test-retest reliability of versions A to B range from .59-.86[116]. Four of the eight subtests will be administered in this study and include:
        - **Visual Elevator.** Participants count up and down floors as they follow a series of visually presented doors and arrows. An accuracy (how many final floor numbers the participant gets correct out of 10) and timing scores (total time taken for the correct items) are

obtained. The reported score is the total time taken divided by the number of switches for the correct items and the scaled score. This subtest measures attentional switching[117].

- **Elevator Counting wth Reversal.** As with the Visual Elevator subtest, participants count up and down as they follow a series of visually presented doors. In this subtest however, the tones are at a fixed speed. The recorded score is the number of correct answers out of 10 and the scaled score. This subtest measures auditory-verbal working memory[117].
- **Telephone Search.** Participants search for symbols while searching for entries on a replicated page of a telephone directory. The recorded score is the time per target score (number of correctly detected symbols divided by the time taken to identify the correctly detected symbols) and the scaled score. This subtest measures selective attention[117].
- **Telephone Search While Counting.** Participants search the telephone directory while counting strings of tones simultaneously. The recorded score is obtained by the combined performance on the Telephone Search and Telephone Search While Counting gives a measure of divided attention, as well as the scaled score. This subtest measures sustained attention[117].

Exploratory measures

Exploratory mesures include a measure of music aptitude and two self-efficacy measures.

Advanced Measures of Music Audiation (AMMA) consists of 30 paired piano melodies that is used to evaluate music aptitude[118]. Test -retesteliability for music majors is r=.89 and non-music majors, r=.83[118].

The General Self-Efficacy[119] is comprised of 23 positive and negative items on a 14-point Likert scale ranging from strongly disagree (1) to strongly agree (14). Internal reliability is established (α=.86-.71). Items are designed to reflect elements of effort, initiation, and persistence as described in Bandura's social cognitive theory.

An adapted version of the Music Performance Self-Efficacy Measure[120] will be used. This questionnaire (α=.97) consists of 24 items to which participants respond by selecting a number (0–100) to reflect strength of agreement to a specific statement regarding musical beliefs (0 = strongly disagree … 100 = strongly agree). The items included are based on Bandura's sociocognitive theory and assess domains: mastery experiences (eight items), vicarious experiences (five items), verbal/social persuasion (six items), and physiological state (five items). An example of an item from this measure is, "I have had positive experiences performing music in the past." This measure has been used in studies with older adults[121]. The adaptation from the original version includes a change in the order of the questions and changes in the wording of 7 questions. More specifically, the question, "I have improved my music performance skills by watching professional musicians, who are similar to me in some way, perform well", was changed to "I have improved my music performance skills by watching professional musicians perform well". The question, "My friends think I am a good performer on my primary instrument", was changed to "My friends think I am a good performer on my piano". The question, "I have had positive experiences performing in large ensembles" was changed to "I have had positive experiences performing in large ensembles (more than 11 performers)". The question, "I have improved my music performance skills by watching other students, who are similar to me in some way, perform well" was changed to "I have improved my music performance skills by watching someone I know perform well (parent, brother, sister, church member, etc,)". The question "I have had positive experiences performing solo, or, in a small ensemble" was changed to "I have had positive experiences performing music solo". The question, "I have had positive

experiences performing music in large ensembles" was changed to " I have had positive experiences peformaing music in a small ensemble (2-10 performers)". Lastly, the question "Performing with my instrument makes me feel good" was changed to "Performing on piano makes me feel good".

6.3.6 Questionnaires

A health questionnaire will be administered at baseline to characterize the sample. Participants are asked whether a doctor or nurse every told them that they any of the following conditions: arthritis, asthma and/or other breathing problems, cancer (other than skin cancer), chronic skin problems, diabetes, heart disease, heart problems (other than heart disease), high cholesterol, hypertention or high blood pressure, mood problems or anxiety, multiple sclerosis, osteoporosis, parkinson's disease, stroke, TIA/mini-stroke within last 18 months, mild cognitive impairment or memory impairment, neuropathy, or any other significant illness.

Participants' expectations about the effects of intervention/control exercises on post-test outcomes will be examined as a potential covariate. During the last week of the intervention stage, the Trainers will administer a modified version of the The Cognitive Training Expectations questionnaire to examine participants' attitudes and expectations about the effects of the intervention received on cognition and everyday function[122]. The NICT questionnaire was modified to refer to "music training" or "piano training", and the "memory" and "reasoning" questions were omitted since we are not assessing these abilities. The quesionnaire explains that cognitive function refers to abilities such as attention, memory, visual perception, information processing, and reasoning; and that cognitive training refers to activities that aim to improve cognitive functions by training within a specific timeframe (i.e., several weeks or months). The questionnaire then asks participants to rate what effects they expect the intervention exercises to have including whether the exercises will result in improved general cognitive function, memory, concentration, distractability, reasoning, multi-tasking, and everday performance. Ratings are on a 7-point Likert scale ranging from completely unsuccessful (1), to no expectations (4), to completely successful (7). Participants also rate the degree to which the intervention was engaging, enjoyable, and challenging on a 7-point scale ranging from very strongly agree (7), to neither agree or disagree (4), to very strongly disagree (1). Finally, participants rate whether or not they were satisfied with the program ranging from extremely satisfied (7), to neither satisfied or dissatisfied (4), to extremely dissatisfied (7) and indicate whether they felt that they had "trained their brain" (yes/no). This questionnaire has been successfully implemented in others' and our prior cognitive intervention studies.

At the baseline visit, a list of prescription medications that the participants' are taking will be recorded. For those who score 20-25 on the MoCA who will be completing a clinical evaluation, the medication list will be provided for the study physician to review. At the clinical evalution visit, the tester will review the list of medications with the participant and document any changes prior to the participant seeing the study physician. The study physician uses this information in the clinical diagnosis or exclusion of MCI.

A questionnaire will be administered at post-test to assess if any health-changes were experienced by the participants during the trial that could affect cognitive or functional performance. Participants are asked if they experienced any significant changes in health conditions such as head injury, chemotherapy or radiation treatment, heart attack or myocardial infarction, changes in vision or hearing, or underwent general anesthesia. They are also asked if they had any new diagnoses of stroke, mini-stroke, TIA, heart disease or congestive heart failure, Parkinson's disease, mild cognitive imairment or memory impairment, Alzheimer's disease, and Multiple Sclerosis.

7 MANAGEMENT OF ADVERSE EVENTS

We will track safety throughout the trial. Participants will be queried as to whether or not they experienced any discomfort at the end of each study visit. See **Reporting of Adverse Events** below for further details. Data and safety monitoring will focus on participant recruitment, data quality, and safety (confidentiality, minimization of risks, and surveillance for adverse events). The study team will meet weekly to discuss the progress of the trial. Details on adverse event monitoring and reporting are provided below.

An adverse event (AE) is defined as any unfavorable and unintended diagnosis, symptom, sign, syndrome or disease which either occurs during the study, having been absent at baseline, or if present at baseline, appears to worsen. Adverse events are to be recorded regardless of their relationship to the study intervention.

Participants may experience fatigue or frustration from completing the intervention exercises. Prolonged sitting activity can result in muscle/joint soreness (e.g., wrist, elbow, fingers, hip flexors, neck, back, buttock), eye strain, muscle cramping or stiffness, or headaches. Any of these symptoms that increase in frequency and last more than 24 hours after participating in intervention exercises will be considered an adverse event. The participant will be advised to seek treatment and not be continued until a physician clears such activity. Serious adverse effects from prior research of the interventions have not been reported.

A serious adverse event (SAE) is generally defined as any untoward medical occurrence that results in death, is life threatening, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity.

8 CRITERIA FOR PARTICIPANT DISCONTINUATION

Participants who undergo any surgical procedures requiring anesthesia, chemotherapy, or radiation during the study period will be excluded. Participants who experience a stroke, TIA, or head injury during the study period will be excluded. Participants who complete any other cognitive intervention not assigned to them during the course of the study will be excluded. In these situations, only data points collected prior to the event will be used in analyses. If plans to undergo such events are known, and 80% of assigned intervention sessions have been offered, an attempt to complete post-test prior to the event will be made.

9 STATISTICAL CONSIDERATIONS

9.1General Design Issues

We will employ a randomized clinical trial design to accomplish the study objectives. Randomization will be stratified based on MCI status. Regardless of randomization or adherence, each eligible and enrolled participant will be encouraged to complete post-test. Participants may be followed longer with positive results and additional funding.

9.2 Primary Objectives

The primary and secondary outcome measures are detailed above.

9.2.1 Primary outcomes

- To examine the efficacy of piano training to improve CAP, cognition, and everyday function among older adults. We will assess the effects of piano training as compared to music listening instruction on CAP, cognition, and everyday function.

CAP includes Time Compressed Speech, Words-in-Noise, Dichotic Digits Test, Dichotic Sentence Identification and Adaptive Tests of Temporal Resolution. All are parametric tests of perception.

Cognition includes Verbal Fluency (phonemic, category, and switching), Digit Coding, and Trail Making test. All are parametric tests of cognitive function.

Everyday function includes the Timed IADL test and Test of Everyday Attention both parametric measures. Both measures will be examined individually.

Principal Component Analyses (PCA) will be performed with CAP, cognitive and functional measures to form composites of outcomes for use in analyses to reduce the number of comparisons.

Hypotheses: The piano training group will outperform the music listening instruction group on measures of CAP, cognition, and everyday function immediately post-intervention.

Secondary Objectives

The secondary objectives are:

- To examine the moderating effects of MCI on piano training efficacy. Moderating effects of MCI status will determine if piano training is effective for older adults with and/or without MCI. We do not have a directional hypothesis for this objective.

- To explore mediators of intervention effects.

  Hypothesis**:** Grounded in theory[8] and prior research[9-11], we hypothesize that piano training will improve CAP and cognitive performance, leading to functional improvements. We expect that enhanced CAP will mediate cognitive gains. We further hypothesize that cognitive gains will mediate functional improvements.

9.3 Sample Size

We expect up to 20% attrition after the initial in-person visit across the time period in which cognitive intervention study protocols will be administered. Our goal is to complete in-person screening visits for up to 450 participants in order to have up to 360 participants complete the study. We can tolerate higher rates of attrition. We require a minimum of 198 eligible participants with *at least* 99 individuals randomized to each condition (piano training, controls) to have adequate statistical power. With stratified randomization, this will result in *at least* 48 persons per MCI/no MCI subgroup.

Statistical power analysis using G*Power[123] in consideration of our pilot data effect sizes as well as attrition was conducted to estimate power. We will control type I error at .05 to achieve at least 80% statistical power.

To power **Specific Aim 1: To examine the efficacy of piano training to improve CAP, cognition, and everyday function among older adults**, a final *n* of 198 participants (n=99 per randomized condition) will achieve 95% power to detect a significant group x time interaction across baseline to immediate post-test, with an alpha of .05. This assumes a small effect size (*f*) of 0.10[124] (Cohen's *f* effect sizes are 0.10- small, 0.25- medium: *f* is roughly equivalent to ½ of *d* when comparing two groups)[124].

To power **Specific Aim 2: To examine the moderating effects of MCI on intervention efficacy,** a final n of 168 will have .95 power at alpha .05 to detect medium effects (*f*=.10) for significant group x time x moderator three way interactions.

9.4 Data Monitoring

To achieve robust and unbiased results, data will undergo a rigorous quality control process to ensure consistency in scoring, coding and accuracy of data entry. Additionally, 20% of all data folders will undergo a random audit every replicate. Also see **Data Management** procedures and **Quality Assurance** detailed below.

9.5Interim Data Analyses

Interim analyses will be conducted in June of 2019, after the pilot, first, and second replicates are completed. We expect that 140 participants will have completed post-test at this time. The purpose will be to confirm feasibility of the training conditions, to assess effect size on cognitive measures from pre- to post- training and adapt number of sessions, if warranted, and to monitor safety.

Interim analyses will be conducted to ensure the safety of the trial. Per NIA, this project does not meet the criteria for requiring a formal Data and Safety Monitoring Plan or Data Safety Monitoring Board because it is not a Phase III clinical trial as defined by NIH. As a behavioral intervention study, this is a phase II efficacy trial[1]. At interim analysis, Dr. Ji will compare the number of adverse events between the two randomized conditions to assess safety. The two groups will also be compared on attrition rates. Any significant differences between the two conditions in adverse events or attrition rates will be further investigated by Dr. Ji. If there is a significant imbalance, confirmed, then Dr.Ji will report the findings to the Clinical Trial Coordinator, Dr. Hudak, who is unblinded. The Clinical Trial Coordinator will work with the IRB to decide if the study should be terminated early.

Interim analyses will be conducted to ensure the feasibility of study intervention adherence. If at interim analyses, less than 75% of participants randomized to an arm complete at least 16 sessions (80% of prescribed 20), the intervention arm will be deemed not feasible. If this occurred for both arms, the study will end.

Finally, interim analyses will examine the potential efficacy of the piano training arm. Effect sizes and confidence intervals for all cognitive outcomes will be calculated for pre- to post- training relative to controls. We expect that the piano training arm will produce an improvement of at least *d*=0.25 from baseline to post-test on at least one of the cognitive measures of verbal fluency (phonemic, category or switching), Trails A or B, or Digit Coding, relative to active controls. If an effect size from pre- to post- training relative to controls of this

magnitude is not observed on at least one of these outcomes, we will adapt study procedures by changing the number of training sessions from 20 to 25. In this instance number of training sessions completed will be examined as a covariate in final analyses.

9.6 Final Data Analyses

Dealing with Score Distribution. We will inspect the distribution of scores to confirm that parametric statistical testing is appropriate. In the event that distributional properties do not support parametric tests, we will compare the results from conventional parametric analyses with (a) results from similar nonparametric testing (e.g., Wilcoxon signed ranks test), (b) results from parametric analyses conducted after the non-normally distributed variables are transformed to improve normality to compare the results with the conventional parametric analyses, and (c) Any illegal values or outliers of >= +/-3*z* will be checked for accuracy with the raw data.

Composite Outcomes. Data reduction in the form of principal component factor analysis (PCA) will be used to derive composite outcomes of CAP,cognition, and everyday function for analyses to reduce the number of comparisons. Baseline data across CAP, cognitive, and everyday function outcomes will be factor analyzed using pairwise deletions. Based on the number of factors derived and factor loadings of the individual measures (.4 or greater), outcome composite scores will be calculated after transforming the baseline data to z scores while taking into account the directional scaling of the items. Post-test scores will be standardized based upon the baseline mean and SD. The baseline factors and loadings will be applied to calculate post-test composite scores.

Preliminary/Descriptive Analyses. Prior to the analyses described below, we will conduct one-way ANOVA or Chi-square analyses as appropriate comparing the randomized groups at baseline on age, race, education, sex, MoCA, GDS, hearing and vision, to ensure that the study randomization procedures were successful. We will also confirm the equivalence of the two randomized conditions by comparing the NICT questionnaire responses. The total across the six items relevant to cognition will be calculated and examined in analyses. If even a marginally significant ($p<.1$) difference is observed for any of these study variables, the pertinent variables will be included as covariates in subsequent analyses.

Specific Aim 1: To examine the efficacy of piano training to improve CAP, cognition, and everyday function among older adults. We will compare changes in performance for piano training relative to control across the two measurement points (baseline, immediate post-test). We will use the PCA-derived composites of CAP and cognition, and everyday functioning as outcomes. We will use repeated measures analyses of variance within the mixed effects models, examining group as the between-subjects factor and time as the within-subjects factor. The group x time interaction will indicate whether the participants randomized to piano training change at differential rates over time relative to those randomized to the control group. This output will be obtained for each of the composite outcomes (CAP, cognition, and functional abilities). For aim 1 analyses and the composite outcomes of CAP, cognition, and everyay function, we will apply Holm-Bonferroni adjustments[125] to reduce Type I error. Specifically, the Holm-Bonferroni procedure uses a step-down process whereby p-values are sorted from lowest to highest, then the lowest p-value is used in it raw version, the second lowest p-value is multiplied by two, the third lowest p-value is multiplied by three and so on until the adjusted

p-value reaches or exceeds the pre-specified threshold for statistical significance, in our case a two-tailed p value of .05.

Intent-To-Treat Approach. We expect that some participants will not complete their intended followup according to the protocol for a variety of reasons, generating missing values. Intent-to-treat (ITT) analysis has become a widely accepted method for dealing with this issue. To maximize available information, we will employ statistical techniques that allow for the inclusion of missing data. Specifically, we will use linear mixed effects models to analyze data. This advanced analytical method allows for the inclusion of participants with missing data, so long as data are missing at random and offers flexibility with regard to the specification of the variance/covariance structure, which often leads to improved fit of the analytical model to the actual data. If there is non-ignorable missing data, we will apply pattern-mixture mixed effects models to assess the sensitivity of our findings with regard to missing data [126].

Specific Aim 2: To examine the moderating effects of MCI on intervention efficacy. We will also test for moderation effects to determine at what stage along the cognitive status continuum—without MCI vs. MCI—intervention is most effective. To do so, we will examine three-way interactions (MCI x training group x time) for the CAP, cognitive, and everyday function composites. When an interaction is significant, we will stratify analyses by MCI status and examine effect sizes. Effect sizes from pre- to post- training relative to controls will be compared for participants with/without MCI. An alpha of .05 will be used for these analyses.

Participants who score 26 or higher on the MoCA will be defined as no MCI (i.e., cognitively normal). Those participants who score a 20-25 on the MoCA will be defined as MCI if they score a 0.5 on the CDR and have performance at or 1 SD below their age, sex, and education norms on any of the following measures: Trails, Verbal Fluency, Digit Coding, Craft Story Recall, Multilingual Naming Test, or Benson Complex Figure. Those with MoCA scores 20-25 who score 0 on the CDR, but have performance at or 1SD below expected for age, sex, and education will be adjudicated prior to analyses. If an individual scores 20-25 on the MoCA, and 0.5 on the CDR, but has performance above expected on all of the aforementioned measures, their case will be adjudicated prior to analyses. The study physicians will adjudicate such cases prior to coding of MCI status for final analyses.

To quantify the effects of piano training further, in addition to examining composite outcomes as detailed above, we will calculate and report effect sizes for pre-to-post change of piano training relative to controls for each individual measure of CAP, cognition, and everyday function for the sample overall as well as stratified by MCI status and among those who were adherent only (80% or more of sessions completed).

Additional Analyses. To address exploratory aim 3, with identification of significant effects from the analyses for aim 1, we will examine behavioral mediators of interventions to determine mechanisms. We hypothesize that improved CAP will mediate cognitive gains. We hypothesize that cognitive gains will mediate functional gains. Formal mediation tests using a bootstrapping technique as outlined by Hayes[127] and Selig and Preacher[128] will be used to examine these hypotheses. The pre-to-post differences will be captured by a latent score[128], whereby the post-test score will serve as the outcome with pre-test score covaried. Results yield estimates and confidence intervals for the indirect and total effects. The procedure allows simultaneous estimation of all three pathways (M on Y; X on M; X on Y, where M=mediator (CAP score), Y=outcome (cognition), and X=predictor (group assignment)) as well as adjustment for covariates (e.g., depression). The use of bootstrapping will ensure that power is not reduced from the main effect analyses[129]. An alpha of .05 will be used for these analyses.

9.7 Sensitivity Analyses

Sensitivity analyses will be conducted to examine if missing data or outliers significantly affected results. We will conduct sensitivity analysis for missing data by using pattern mixture models. The pattern of results will be compared to ITT analyses. Results will also be conducted after any outliers, defined as observations with values of greater than +3 or smaller than -3 z score units, are removed and compared to primary results.

9.8 Exploratory Subgroup Analyses

Pursuant to NIH guidelines on examining sex as a biological variable, we will examine if sex moderates any significant intervention effects identified in aim 1 or aim 2 analyses by examining sex x group x time interactions. Effect sizes will be reported by sex if there are any significant differences (alpha .05).

Relevant to aim 1 and 2 analyses, exploratory analyses will further examine effect size for the pre-to post-change of piano training relative to music listening instruction among those who were adherent (i.e., who completed 80% or more of assigned number of sessions) in both conditions. Analyses detailed above for specific aims 1 and 2 will be repeated among this subsample. Results and effects sizes among the adherent sample will be compared to the ITT derived results and effect sizes.

Relevant to aim 1 and 2 analyses, we will further examine if effects were moderated by a better ear pure-tone average (PTA). In these analyses we will examine if there are significant training group x time x hearing interactions on CAP, cognition, or everyday function, using an alpha of .05.

As it is possible that training will be differentially effective for those with CAP disorders (CAPD), we will identify a subgroup of participants (regardless of MCI status) who could be diagnosed with CAPD. Diagnosis of (C)APD requires performance deficits on the order of at least two standard deviations below the mean on two or more of the following tests[39, 130]: temporal processing (TCS, ATTR), degraded speech understanding (WIN), or binaural processing (DDT, DST). Analyses for the cognition and everyday function outcomes will be repeated with groups defined as with/without CAPD as an independent variable. We will examine if there are significant training group x time x ADP status interactions using an alpha of .05.

We may also explore moderation effects of covariates such as MCI subtype (amnestic, non-amnestic, multiple domain) to determine who is most likely to benefit from the intervention among participants with MCI. To do so, we will examine three-way interactions (covariate x training group x time) stratified by MCI status, if MCI was a significant moderator in results of analyses for aim 2 described above. Effect sizes will be examined.

The self-efficacy measures (general self-efficacy and music performance self-efficacy) will be examined as exploratory outcomes by comparing the two randomized conditions across the two measurement points (before/after training). Analyses will include ordinal logistic regression analyses across levels of self-eficacy conducted within the generalized estimating equations (GEE) statistical framework to account for repeated measurements. Significant improvements on either measures demonstrated in this analyses will be followed by exploratory mediation analyses to examine if improvements in self-efficacy can account for any observed training

effects on CAP, cognition, or everyday function. Thus, we will examine an alternative hypothesis that training gains can be attributed to improved self-efficay. An alpha of .05 will be used.

To explore if initial music aptitude affects performance within each of the training conditions, AMMA at baseline will be examined as a covariate (i.e., modifier) of Basic Piano Measure performance (BPM) from the first to the final day of training among those in the piano training condition. Similarly, baseline AMMA will be examined as a covariate on the Music Listening Measure (MLM) across the same two time points within the control group. Again, ordinal logistic regression with GEE will be used given the ordinal scaling of the outcomes. An alpha of .05 will be used.

10 DATA COLLECTION, SITE MONITORING, AND ADVERSE EXPERIENCE REPORTING

10.1 Records to Be Kept

A visit control sheet (VCS) will be the only document indicating the participants name and subject id. This data form will be kept in a locked filing cabinet separate from the study data. This data form will be accessible to the Testers and Trainers as well as PIs, only.

Signed consent forms will be kept in locked filing cabinets separate from the data forms and will not contain the subject id number.

We will keep all completed telephone screening interviews to track the number of screenings completed, number scheduled, and reasons ineligible. We will keep the data forms completed and intervention records, which will contain subject id only, in a locked filing cabinet.

All documents will be kept for at least 5 years post-study close out (or longer if required by site IRB or NIH or other governing body).

The study database will not include participant names or other identifying information. When the study data are released for sharing, the data will be completely de-identified with any potentially identifying information (e.g., date of birth) removed.

10.2 Data Management

10.2.1 Clinical site responsibilities in data collection and management.

Each study participant will have a folder with participant id for all assessments completed for each study visit and for the intervention phase. Every data form (page) will include the participant id and date the form was administered.

Original versions of the data forms arranged by study visit will be kept in a locked filing cabinet in the lab noted with an electronic file location in the footer. Data forms will only be copied from these originals only. Each testing visit will include a detailed checklist of all measures to be administered and procedures to be followed, which will be initialed by the tester upon completion.

Immediately after completing a participant visit, the tester will score the assessments and check all forms for completion, and within 21 days a different tester should use a detailed checklist to review all of the data forms for accuracy and completion. Any unclear or missing items are queried to the Tester or study participant as appropriate. This quality control process helps to avoid missing data. This process will be repeated as necessary prior to data entry until the checklist is completed. This rescoring process is the first step of our data quality control process.

At the end of each testing visit and intervention phase, every data form will be reviewed in a five-step quality control audit (scoring, rescoring, entry, checking, reconciling) with each subsequent step performed by a different Research Assistant.

The data will be entered into a custom-designed MS Access database that uses MS SQL Server as the back end relational database and completely integrates study management (e.g., create forms, collect data, generate reports, prepare for analysis). The database is designed to ensure access control, audit control, data integrity, user authentication, and transmission security. The data quality control and custom database function to minimize missing data and ensure data integrity.

Only individuals trained by the PI or Study Coordinator to use the MS Access database will enter data. Reports reflecting entered data will be printed for each testing visit of each participant upon data entry. The data report will be checked for accuracy with the raw data by a trained staff member other than who entered the data. Errors or omissions will be returned for reconciliation and the process will be repeated until the entered data are deemed error free. Additionally, a random audit of 20% of folders entered will be completed each replicate.

11 Quality Assurance

The study protocol including a priori hypotheses and data analytic plan will be published online within 90 days of commencement of data collection. Any changes to the protocol or analytic plans will be documented with date of change and justification.

All study staff who recruit, interview, or assess study participants will be required to attend a weekly staff meeting with the Principal Investigator throughout the data collection phase of the study. Recruitment goals and procedures will be reviewed. Problems or issues will be discussed and resolved. Detailed minutes of the meetings proceedings will be kept for review by any personnel who are on vacation or ill and miss a meeting. If staff change, new personnel will review all previous study meeting minutes prior to working on the project. The study coordinator will keep these records from each staff meeting.

The study investigators will meet about monthly during the first quarter of the study, will meet at interim analyses, and at least annually to review/monitor any adverse events, if they occur. The study investigators may also meet monthly in the final quarter of the study to complete data analysis, interpretation, and publication.

All study staff will read, be instructed in, and may be asked to pass a certification exam on the study protocol before working on assigned study activities. All study staff will be required to complete IRB Human Subjects training prior to working on the project with re-certification completed as required by the NIH or IRB. This will be monitored and implemeneted by the study coordinator.

Telephone screeners will be required to complete two practice inteviews with the first actual study interview supervised and evaluated by the Principal Investigator, Study Coordinator, or trained and certified telephone screener. The Study coordinator or other site personnel appointed to do will audit the first telephone screening interview completed and will continue to do so until the questionnaire and final outcome are completed without error.

The Testers, who will administer the baseline, clinical and post-test visits, will be required to practice administration of all study measures per the study protocol with three persons. The testing must be completed with at least two individuals who meet the study age requirements. The third practice visit will be supervised by the Principal Investigator, Clinical Coordinator, or a trained and certified tester (more practice may be deemed necessary by the observer). The first actual study assessment visit will also be supervised and evaluated by one of these study personnel. To further achieve internal validity and inter-rater reliability, the testers will observe and evaluate each other at least once per replicate and audit each other's testing records. Participant data forms should be audited by a different tester within 21 days of completion per a detailed rescoring checklist. Any missing, incomplete, or unclear items will be queried to the tester or participant as appropriate. The Study Coordinator will implement and monitor this process. Additional training and certification is required to administer the Clinical Dementia Rating Scale.

A detailed intervention protocol will be developed for Trainers. The Trainers may be required to pass a certification test of the intervention protocol prior to interacting with study participants during the intervention phase. The Trainers will be required to complete at least 5 hours of training activities that they will administer as a part of their training, prior to administering the intervention exercises to actual study participants. They will complete an initial mock training session under the observation of the PI, Study Coordinator, or another trained Research Assistant Trainer prior to conducting an actual study intervention visit. To further achieve internal validity and intervention fidelity, the trainers will observe and evaluate each other at least once per replicate and audit each other's intervention records. The first actual intervention visit conducted by the trainer will also be supervised and evaluated by the study coordinator or a certified trainer. Participant training data forms will be audited by a different trainer within 21 days of completion per a detailed rescoring checklist. Any missing, incomplete, or unclear items will be queried to the trainer or participant as appropriate. The Study Coordinator will implement and monitor this process.

The Study coordinator or other site personnel appointed to do will audit the first testing/intervention folder completed and rescored by each tester/trainer, and will continue to do so until the procedure is completed without error. Random quality control audits will be performed on 20% of the participant folders each replicate.

Study documents and pertinent records will be available for inspection by monitoring authorities (e.g., NIH or IRB) as requested.

11.1 Adverse Experience Reporting

Study staff will monitor participants and be alert to any adverse events at each study.. Testers and Trainers will ask participants about any adverse events at the end of each testing and intervention visit. Furthermore, the participants will be interviewed about any such events throughout their study participation at the post-test visit. Study staff will be alert to any volunteered adverse events. All such events will be documented on a standardized form. At all in-person contacts and at any other applicable time throughout the study, any potential adverse events reported by a subject or observed by the research staff will be recorded and subsequently evaluated by the Principal Investigator for its relation to the study research procedures and whether or not any corrective action need be taken. It is unlikely and not expected that any adverse incident will result from implementation of this study protocol. However, any and all adverse events that may occur will be recorded on the appropriate forms, and reported to the governing IRB, as applicable.

Any adverse events will be reported to the site IRB within 72 hours of occurrence or detection by PI Edwards. In the unlikely event of a serious adverse event this report will be made within 24 hours and reported also to the NIH Project Officer. In the unlikely case of study-related serious adverse events, an interim analysis will be performed in order to determine whether a change in the risk/benefit ratio has occurred and reviewed by the statistician Dr. Ji. If so, this change will be brought to the attention of the USF IRB for review, current and future participants will be notified of this change, and stopping rules will be considered.

11.2 Follow-up for Adverse Events

Research Assistants will monitor participants for any adverse events at study visits. At the post-test study visit, participants will be directly interviewed about any possible adverse events that occurred since their initial visit. A study physician and/or Dr. Edwards will follow-up with the participant for any adverse events. The participant will be referred to proper treatment.

11.3 Safety Monitoring

The PI and the Study Physicians will be responsible for ensuring participants' safety.

Dr. Ji will perform analyses to monitor participant safety, evaluate the progress of the study, to review procedures for maintaining the confidentiality of data, the quality of data collection, management, and analyses.

All OHRP and local IRB requirements for reporting adverse events will be followed.

11.4 Confidentiality

All evaluation forms, reports, recordings, and other records that leave the site will be identified only by the participant id number to maintain subject confidentiality. All records will be kept in a locked file cabinet. All computer entry of data forms will be done using participant ids only and without participant names. Clinical information will not be released without written permission of the participant, except as necessary for monitoring by IRB, NIH, or the OHRP

11.5 Study Modification/Discontinuation

The study may be modified or discontinued at any time by the IRB, the NIH, the OHRP, or other government agencies as part of their duties to ensure that research subjects are protected.

Study Modifications

Date Implemented	Modification	Rationale
04/23/2018	For the pilot replicate, the FAS, Animals/ Boys' Names, and Fruits/ Furniture version will be administered at baseline and	The music listening condition involves a lesson on musical intstruments. Thus, we decided that using this version of verbal

	also at post test.	fluency category switching at post would unfairly bias the active control condition. We did not realize this until during the training phase of the pilot replicate. We changed the order for furture replicates such that Vegetables/Musical Instruments version is administered at baseline and the Fruits/Furniture version will be administered at post-test.
07/16/2018	Page 12, 5.2 Exclusion Criteria. We clarified that individuals who have previously participated in the USF Music Research and Testing Lab studies are excluded.	This was inadvertently omitted from the protocol version 1. All individuals screened for the study were excluded if they reported prior participation in the Music Research and Testing Lab. The rationale is such individuals would have completed similar or the same measures as administered in this study and may have received music training.
07/16/2018	Page 15, 5.7 Interventions Administration and Duration. We omitted the sentence that indicated a textbook will be provided and weekly homework will be assigned.	The study investigators met to discuss the study progress and any modifications needed since completion of the pilot replicate. The trainers and a Co-I noted that it was difficult to make the homework assignments and engagment equivalent across the two conditions. It was decided that we will omit the homework for replicate 1 and the remainder of the study. Participants will only access the textbooks during the in-lab training sessions. There is no evidence that homework is required for piano training to be efficacious. Ensuring equivalency of the two conditions is the rationale.
07/16/2018	Page 12, 5.2 Exclusion criteria. We added, “which interferes	The PI and study coordinator felt the exclusion criteria of any

	with ability to use a keyboard" to exclusion criteria regarding numbness or tingling.	numbness or tingling in any finger was too conservative and that individuals were being excluded that would be able to successfully participate in the study. We decided to add this phrase in an attempt to only exclude those who would not be able to successfully complete the intervention.
07/16/2018	Pages 13-14, 5.3 Enrollment Procedures. Those who score a 26 or better on the MoCA will be enrolled in the study (not 27). Those potentially eligible who score between 20 to 25 (not 26) on the MoCA who were not referred by the MDC or study physician will also complete a clinical evaluation visit.	The PI and Study Coordinator felt the criteria for clinical evaluation were too stringent. The study investigators decided to use the traditional cut-point of 26 for MoCA to determine who is eligible and who requires further evaluation.
07/30/2018	Page 14, 5.3 Enrollment Procedures. We clarify that clinically significant abnormalities are shared with the participants who are encouraged to seek treatment. We added the phrase, "that would likely interfere with their ability to benefit from intervention" regarding any significant clinical abnormalities.	This has been our procedure beginning in replicate 1. The goal is to be inclusive and to allow those who could potentially benefit from intervention to continue participation in the study.

12 REFERENCES

1. Onken LS, Carroll KM, Shoham V, Cuthbert BN, Riddle M Reenvisioning Clinical Science: Unifying the discipline to improve the public health Clinical Psychological Science 2014;2(1):22-34 doi: 10.1177/2167702613497932258216582. Gates GA, Gibbons LE, McCurry SM, Crane PK, Feeney PK, Larson EB Executive dysfunction and presbycusis in older persons with and without memory loss and dementia Cognitive and Behavioral Neurology 2010;23:218-23 doi: 10.1097/WNN.0b013e3181d748d73. Gates GA, Anderson ML, Feeney MP, McCurry SM, Larson EB Central auditory dysfunction in older persons with memory impairment or Alzheimer dementia Archives of Otolaryngology -- Head and Neck Surgery 2008;134(7):771-72008/07/23 doi: 10.1001/archotol.134.7.771186451304. Gold M, Lightfoot LA, Chisolm T Hearing loss in a memory disorders clinic: A specially vulnerable population Archives of Neurology 1996;53:922-85. Albers MW, Gilmore GC, Kaye JA, Murphy C, Wingfield A, Bennett DA, Boxer AL, Buchman AS, Cruickshanks K, Devanand DP, Duffy CJ, Gall CM, Gates GA, Granholm TH, Holtzer R, Hyman B, Lin F, McKee JQ, Trojanowski JQ, Verghese J, Wilson DA, Xu S, Zhang LI At the interface of sensory and motor dysfunctions and Alzheimer's disease Alzheimer's & Dementia: The Journal of the Alzheimer's Association 2015;11:70-98 doi: 10.1016/j.jalz.2014.04.5146. Gates GA, Beiser A, Rees T, D'Agostino R, Wolf P Central auditory dysfunction may precede the onset of clinical dementia in people with probable Alzheimer's disease Journal of the American Geriatrics Society 2002;50:482-8 doi: 10.1046/j.1532-5415.2002.50114.x7. Gates GA, Cobb J, Linn R, Rees T, Wolf P, D'Agostino R Central auditory dysfunction, cognitive dysfunction, and dementia in older people Archives of Otolaryngology -- Head and Neck Surgery 1996;122:161-7 doi: 10.1001/archotol.1996.018901400470108. Schneider BA, Pichora-Fuller MK. Implications of perceptual deterioration for cognitive aging research. In: Craik FIM, Salthouse TA, editors. The Handbook of Aging and Cognition. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 2000. p. 155-220.9. Nikjeh D, Lister J, Frisch S Hearing of note: An electrophysiologic and psychoacoustic comparison of pitch discrimination between vocal and instrumental musicians Psychophysiology 2008;45:994-100710. Nikjeh D, Lister J, Frisch S Pre-attentive cortical evoked responses to pure tones, music, and speech: Influence of music training Ear and Hearing 2009;30:432-4611. Alain C, Bidelman G Musical Training Orchestrates Coordinated Neuroplasticity in Auditory Brainstem and Cortex to Counteract Age-Related Declines in Categorical Vowel Perception Journal of Neuroscience 2015;35(3):1240-912. Morris JC, Storandt M, Miller JP, McKeel DW, Price JL, Rubin EH, Berg B Mild cognitive impairment represents early-stage Alzheimer's disease Archives of Neurology 2001;58:397-405 doi: 10.1001/archneur.58.3.397.13. Corner L, Bond J Being at risk of dementia: Fears and anxieties of older adults Journal of Aging studies 2004;18:143-55 doi: 10.1016/j.jaging.2004.01.00714. Seshadri S, Wolf PA Lifetime risk of stroke and dementia: Current concepts, and estimates from the Framingham Study Lancet Neurology 2007;6(12):1106-14 doi: 10.1016/S1474-4422(07)70291-01803170715. Sands LP, Yaffe K, Lui LY, Stewart A, Eng C, Covinsky K The effects of acute illness on ADL decline over 1 year in frail older adults with and without cognitive impairment Journals of Gerontology Series A, Biological Sciences and Medical Sciences 2002;57:M449-M54

16. Kirschstein R. Disease-specific estimates of direct and indirect costs of illness and NIH support: Fiscal year 2000 update: National Institute of Health; 2000. Available from: http://ospp.od.nih.gov/ecostudies/COIreportweb.htm.
17. Hurd MD, Martorell P, Delavande A, Mullen KJ, Langa KM Monetary costs of dementia in the United States New England Journal of Medicine 2013;368(14):1326-34 doi: 10.1056/NEJMsa120462923550670
18. Hobbs FB. Population of the United States 2008. Available from: http://www.census.gov/population/www/pop-profile/elderpop.html.
19. Ferri CP, Prince M, Brayne C, Brodaty H, Fratiglioni L, Ganguli M, Hall K, Hasegawa K, Hendrie H, Huang Y, Jorm A, Mathers C, Menezes PR, Rimmer E, Scazufca M Global prevalence of dementia: A Delphi consensus study Lancet 2005;366:2112-7 doi: 10.1016/S0140-6736(05)67889-0
20. Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM Forecasting the global burden of Alzheimer's disease Alzheimer's & Dementia 2007;3:181-91 doi: 10.1016/j.jalz.2007.04.381
21. DeCarli C Mild cognitive impairment: Prevalence, prognosis, aetiology, and treatment The Lancet Neurology 2003;2:15-21 doi: 10.1016/S1474-4422(03)00262-X
22. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos E, Kokmen E Mild cognitive impairment: Clinical characterization and outcome Archives of Neurology 1999;56:303-8 doi: 10.1001/archneur.56.3.303
23. Petersen RC Mild cognitive impairment as a diagnostic entity Journal of Internal Medicine 2004;256(3):183-94 doi: 10.1111/j.1365-2796.2004.01388.x
24. Reuter-Lorenz PA, Park DC How does it STAC up? Revisiting the scaffolding theory of aging and cognition Neuropsychology Review 2014;24(3):355-70 doi: 10.1007/s11065-014-9270-925143069
25. Belleville S Cognitive training for persons with mild cognitive impairment International Psychogeriatrics 2008;20(1):57-66 doi: 10.1017/S104161020700631X
26. Stott J, Spector A A review of the effectiveness of memory interventions in mild cognitive impairment (MCI) International Psychogeriatrics 2010;23:526-38 doi: 10.1017/S1041610210001973
27. American Speech Language Hearing Association. Guidelines for manual pure-tone threshold audiometry [guidelines]. 2005.
28. Gates GA, Gibbons LE, McCurry SM, Crane PK, Feeney MP, Larson EB Executive dysfunction and presbycusis in older persons with and without memory loss and dementia Cognitive and Behavioral Neurology 2010;23(4):218-23 doi: 10.1097/WNN.0b013e3181d748d721150347
29. Li KZH, Lindenberger U Relations between aging sensory/sensorimotor and cognitive functions Neuroscience and Biobehavioral Reviews 2002;26(7):777-83 doi: 10.1016/S0149-7634(02)00073-8
30. Fabrigoule C, Rouch I, Taberly A, Letenneur L, Commenges D, Mazaux J, Orgogozo J, Dartigues JF Cognitive process in preclinical phase of dementia Brain 1998;121(1):135-41
31. Grober E, Hall CB, Lipton RB, Zonderman AB, Resnick SM, Kawas C Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease Journal of the International Neuropsychological Society 2008;14(2):266-78 doi: 10.1017/S1355617708080302 18282324
32. Belleville S, Chertkow H, Gauthier SUS Working memory and control of attention in persons with Alzheimer's disease and mild cognitive impairmentUSUS NeuropsychologyUS 2007;21(4):458-69US Epub US doi: 10.1037/0894-4105.21.4.458US PubMed PMID: US
33. Stopford CL, Thompson JC, Neary D, Richardson AM, Snowden JS Working memory, attention, and executive function in Alzheimer's disease and frontotemporal dementia Cortex 2010;48(4)2011/01/18 doi: 10.1016/j.cortex.2010.12.002 21237452
34. Gates GA, Beiser A, Rees TS, D'Agostino RB, Wolf PA Central auditory dysfunction may precede the onset of clinical dementia in people with probable Alzheimer's disease Journal of the American Geriatrics Society 2002;50(3):482-8 11943044

35. Gates GA, Anderson ML, McCurry SM, Feeney MP, Larson EB Central auditory dysfunction as a harbinger of Alzheimer dementia Arch Otolaryngol Head Neck Surg 2011;137(4):390-5 doi: 10.1001/archoto.2011.2821502479
36. Humes LE, Dubno JR, Gordon-Salant S, Lister JJ, Cacace AT, Cruickshanks KJ, Gates GA, Wilson RH, Wingfield A Central presbycusis: A review and evaluation of the evidence Journal of the American Academy of Audiology 2012;23(8):635-66 doi: 10.3766/jaaa.23.8.522967738
37. Edwards JD, Lister JJ, Elias MN, Tetlow AM, Sardina AL, Sadeq NA, Brandino AD, Harrison Bush AL Auditory processing of older adults with probable mild cognitive impairment Journal of Speech, Language, and Hearing Research 2016
38. Idrizbegovic E, Hederstierna C, Dahlquist M, Kampfe Nordstrom C, Jelic V, Rosenhall U Central auditory function in early Alzheimer's disease and in mild cognitive impairment Age and Ageing 2011;40(2):249-54 doi: 10.1093/ageing/afq16821233090
39. Association AS-L-H. Guidelines for manual pure-tone threshold audiometry [guidelines]. 2005.
40. American Speech-Language-Hearing Association. Central auditory processing: Current status of status of research and implications for clinical practice. 1996.
41. Pichora-Fuller MK, Singh G Effects of age on auditory and cognitive processing: Implications for hearing aid fitting and audiologic rehabilitation Trends in Amplification 2006;10:29-59 doi: 10.1177/108471380601000103
42. Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N Reversal of age-related neural timing delays with training Proceedings of the National Academy of Sciences of the United States of America 2013;110(11):4357-62 doi: 10.1073/pnas.121355511023401541
43. Pichora-Fuller MK, Schneider BA, Daneman M How young and old adults listen to and remember speech in noise Journal of the Acoustical Society of America 1995;97:593-608
44. Rabbitt PM Channel-capacity, intelligibility and immediate memory Quarterly Journal of Experimental Psychology 1968;20(3):241-8 doi: 10.1080/14640746808400158568376 3
45. Tun PA, McCoy S, Wingfield A Aging, hearing acuity, and the attentional costs of effortful listening Psychology and Aging 2009;24(3):761-6
46. Grimes AM, Grady CL, Foster NL, Sunderland T, Patronas NJ Central auditory function in Alzheimer's disease Neurology 1985;35(3):352-8 3871922
47. Jäncke L, Shah NJ Does dichotic listening probe temporal lobe functions? Neurology 2002;58(5):736-43
48. Duchek JM, Balota DA Failure to control prepotent pathways in early stage dementia of the Alzheimer's type: Evidence from dichotic listening Neuropsychology 2005;19(5):687-95 doi: 10.1037/0894-4105.19.5.68716187887
49. Kurylo DD, Corkin S, Allard T, Zatorre RJ, Growdon JH Auditory function in Alzheimer's disease Neurology 1993;43(10):1893-9 8413944
50. Sinha UK, Hollen KM, Rodriguez R, Miller CA Auditory system degeneration in Alzheimer's disease Neurology 1993;43(4):779-85 8469340
51. De Santia S, de Leona MJ, Rusinekb H, Convita A, Tarshisha CY, Rochea A, Tsuic WH, Kandila E, Boppanna M, Daisleya K, Wangd GJ, Schlyere D, Folere J, al. e Hippocampal formation glucose metabolism and volume losses in MCI and AD Neurobiology of Aging 2001;22(4): 529–39 doi: 10.1016/S0197-4580(01)00230-5
52. Wingfield A, Peelle J The effects of hearing loss on neural processing and plasticity Frontiers in Systems Neuroscience 2015;9:1-4
53. Rauschecker JP, Scott SK Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing Nature Neuroscience 2009;12:718-24 doi: 10.1038/nn.2331

54. Davis MH, Johnsrude I Hierarchical processing in spoken language comprehension The Journal of Neuroscience 2003;23(8):3423-31
55. Eckert MA, Cute SL, Vaden KI, Jr., Kuchinsky SE, Dubno JR Auditory cortex signs of age-related hearing loss Journal of the Association for Research in Otolaryngology 2012;13(5):703-13 doi: 10.1007/s10162-012-0332-522618352
56. Wong PC, Perrachione TK, Parrish TB Neural characteristics of successful and less successful speech and word learning in adults Human Brain Mapping 2007;28(10):995-1006 doi: 10.1002/hbm.2033017133399
57. Gates GA, Cobb JL, Linn RT, Rees T, Wolf PA, D'Agostino RB Central auditory dysfunction, cognitive dysfunction, and dementia in older people Archives of Otolaryngology -- Head and Neck Surgery 1996;122(2):161-7 8630210
58. Anderson S, Kraus N Auditory training: Evidence for neural plasticity in older adults Perspectives in Hearing and Hearing Disorders: Research and Diagnostics 2013;17:37-57 doi: 10.1044/hhd17.1.3725485037
59. Zendel BR, Alain C Musicians experience less age-related decline in central auditory processing Psychology and Aging 2012;27(2):410-7 doi: 10.1037/a002481621910546
60. Mishra SK, Panda MR, Herbert C Enhanced auditory temporal gap detection in listeners with musical training Journal of the Acoustical Society of America 2014;136(2):EL173-EL8
61. Kraus N, Chandrasekaran B Music training for the development of auditory skills Nature Reviews Neuroscience 2010;11:599-605 doi: 10.1038/nrn2882
62. Humes LE, Busey TA, Craig J, Kewley-Port D Are age-related changes in cognitive function driven by age-related changes in sensory processing? Attention, Perception, & Psychophysics 2013;75(3):508-24 doi: 10.3758/s13414-012-0406-923254452
63. Mahncke HW, Bronstone A, Merzenich M Brain plasticity and functional losses in the aged: Scientific bases for a novel intervention Progress in Brain Research 2006;157:81-109
64. Edwards JD, Clark D, Xu H, Unverzagt FW Computerized process-based training reduces dementia prevalence over ten years (in preparation)
65. Park DC, Lodi-Smith J, Drew L, Haber S, Hebrank A, Bischof GN, Aamodt W The impact of sustained engagement on cognitive function in older adults: The Synapse Project Psychological Science 2014;25(1):103-12 doi: Doi 10.1177/0956797613499592WOS:000329486300011
66. Verghese J, Lipton RB, Katz MJ, Hall CB, Derby CA, Kulansky G, Ambrose A, Sliwinski M, Buschke H Leisure activities and risk of dementia in the elderly New England Journal of Medicine 2003;348:2508-16 doi: 10.1056/NEJMoa022252
67. Seinfeld S, Figueroa H, Ortiz-Gil J, Sanchez-Vives MV Effects of music learning and piano practice on cognitive function, mood, and quality of life in older adults Frontiers in Psychology 2013;4 doi: 10.3389/psyg.2013.00810
68. Paquette S, Goulet GM Lifetime benefits of musical training Frontiers in Neuroscience 2014;8:1-2 doi: 10.3389/fnins.2014.00089
69. Alain C, Zendel BR, Hutka S, Bidelman GM Turning down the noise: The benefit of musical training on the aging auditory brain Hearing Research 2014;308:162-73 doi: 10.1016/j.heares.2013.06.00823831039
70. Oechslin MS, Descloux C, Croquelois A, Chanal J, Van De Ville D, Lazeyras F, James CE Hippocampal volume predicts fluid intelligence in musically trained people Hippocampus 2013;23(7):552-8 doi: 10.1002/hipo.2212023519979
71. Wan CY, Schlaug G Music making as a tool for promoting brain plasticity across the lifespan. Neuroscientist 2010;16:566-77 doi: 10.1177/1073858410377805
72. Bidelman GM, Weiss MW, Moreno S, Alain C Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians European Journal of Neuroscience 2014;40(4):2662-73 doi: 10.1111/ejn.1262724890664

73. James CE, Michel CM, Britz J, Vuilleumier P, Hauert CA Rhythm evokes action: Early processing of metric deviances in expressive music by experts and laymen revealed by ERP source imaging Human Brain Mapping 2012;33(12):2751-67 doi: 10.1002/hbm.2139721932257
74. Strait DL, Chan K, Ashley R, Kraus N Specialization among the specialized: Auditory brainstem function is tuned in to timbre Cortex 2012;48(3):360-2 doi: 10.1016/j.cortex.2011.03.01521536264
75. O'Brien JL, Nikjeh D, Lister J Interaction of musicianship and aging: A comparison of cortical auditory evoked potentials Behavioural Neurology 2015;2015:1-12 doi: dx.doi.org/10.1155/2015/545917
76. Parbery-Clark A, Strait DL, Hittner E, Kraus N Musical training enhances neural processing of binaural sounds The Journal of Neuroscience 2013;33(42):16741-7 doi: 10.1523/JNEUROSCI.5700-12.201324133275
77. Seidler RD Differential effects of age on sequence learning and sensorimotor adaptation Brain Research Bulletin 2006;70(4-6):337-46 doi: 10.1016/j.brainresbull.2006.06.00817027769
78. Seidler RD, Bernard JA, Burutolu TB, Fling BW, Gordon MT, Gwin JT, Kwak Y, Lipps DB Motor control and aging: links to age-related brain structural, functional, and biochemical effects Neuroscience and Biobehavioral Reviews 2010;34(5):721-33 doi: 10.1016/j.neubiorev.2009.10.00519850077
79. Han Y, Yang H, Lv YT, Zhu CZ, He Y, Tang HH, Gong QY, Luo YJ, Zang YF, Dong Q Gray matter density and white matter integrity in pianists' brain: A combined structural and diffusion tensor MRI study Neuroscience Letters 2009;459(1):3-6 doi: 10.1016/j.neulet.2008.07.05618672026
80. Bangert M, Schlaug G Specialization within the specialized in features of external human brain morphology European Journal of Neuroscience 2006;24:1832-4
81. Bugos JA The benefits of music instruction on processing speed, verbal fluency, and cognitive control in aging Music Education Research International 2010;4(1-9)
82. Bugos JA, Perlstein WM, McCrae CS, Brophy TS, Bedenbaugh PH Individualized piano instruction enhances executive functioning and working memory in older adults Aging and Mental Health 2007;11(4):464-71 doi: 10.1080/13607860601086504176128 11
83. Bernard JA, Seidler RD Relationships between regional cerebellar volume and sensorimotor and cognitive function in young and older adults Cerebellum 2013;12(5):1-30 doi: 10.1007/s12311-013-0481-z
84. Miendlarzewska EA, Trost WJ How musical training affects cognitive development: Rhythm, reward, and other modulating variables. Frontiers in Neuroscience 2014;7(1-18) doi: 10.3389/fnins.2013.00279
85. Seppanen M, Pesonen AK, Tervaniemi M Music training enhances the rapid plasticity of P3a/P3b event-related brain potentials for unattended and attended target sounds Attention, Perception, and Psychophysiology 2012;74(3):600-12 doi: 10.3758/s13414-011-0257-922222306
86. Hanna-Pladdy B, Gajewski B Recent and past musical activity predicts cognitive aging variability: Direct comparison with general lifestyle activities Frontiers in Human Neuroscience 2012;6:1-11
87. Hanna-Pladdy B, MacKay A The relation between instrumental musical activity and cognitive aging Neuropsychology 2011;25(3):378-86 doi: http://dx.doi.org/10.1037/a0021895
88. Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment Journal of the American Geriatrics Society 2005;53:695–9
89. Good-Lite. Sloan letter near vision card: Good-Lite; 2010 [October 21, 2010]. Available from: www.good-lite.com.
90. Conradsson M, Rosendahl E, Littbrand H, Gustafson Y, Olofsson B, Lovheim H Usefulness of the Geriatric Depression Scale 15-item version among very old people with and without cognitive impairment Aging & Mental Health 2013;17(5):638-45 doi: 10.1080/13607863.2012.75823123339600
91. Friedman B, Heisel MJ, Delavan RL Psychometric properties of the 15-item geriatric depression scale in functionally impaired, cognitively intact, community-dwelling elderly primary care patients Journal of the American Geriatrics Society 2005;53(9):1570-6 doi: 10.1111/j.1532-5415.2005.53461.x16137289

92. Bugos JA, Groner A The effects of instrumental training on non-verbal reasoning in eighth-grade students Research Perspectives in Music Education 2009;14:14-9
93. Wilson RH Development and use of auditory compact discs in auditory evaluation Journal of Rehabilitation Research and Development 1993;30(3):342-51
94. Wilson RH, Salomon DL, Sperry JL, Bornstein SP Effects of time compression and time compression plus reverberation on the intelligibility of Northwestern University Auditory Test No.6 Journal of the American Academy of Audiology 1994;5:269-77
95. Konkle DF, Beasley DS, Bess FH Intelligibility of time-altered speech in relation to chronological aging Journal of Speech, Language, and Hearing Research 1977;20(1):108-15 doi: 10.1044/jshr.2001.108
96. Beasley DS, Bratt GW, Rintelmann WF Intelligibility of time-compressed sentential stimuli Journal of Speech, Language, and Hearing Research 1980;23:722-31 doi: 10.1044/jshr.2304.722
97. Wilson RH Development of a speech in multitalker babble paradigm to assess word-recognition performance Journal of the American Academy of Audiology 2003;14:453-70
98. Finney DJ. Statistical method in biological assay. 3rd ed. London: Charles Griffin & Co. Ltd; 1964.
99. Wilson RH, Morgan DE, Dirks DD A proposed SRT procedure and its statistical precedent Journal of Speech and Hearing Disorders 1973;38:184-91
100. Wilson RH, McArdle R Intra- and inter-session test, retest reliability of the Words-in-Noise (WIN) Test Journal of the American Academy of Audiology 2007;18(10):813-25
101. Wilson RH, McArdle RA, Smith SL An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. Journal of Speech, Language, and Hearing Research 2007;4:844
102. Strouse A, Hall JW, 3rd Test-retest reliability of a dichotic digits test for assessing central auditory function in Alzheimer's disease Audiology 1995;34(2):85-90
103. Speaks C, Jerger J, Jerger S Performance-intensity characteristics of synthetic sentences. Journal of Speech, Language, and Hearing Research Journal of Speech, Language, and Hearing Research 1966;9(2):305-12 doi: 10.1044/jshr.0902.305
104. Gates GA, Anderson ML, McCurry SM, Feeney MP, Larson EB Central auditory dysfunction as a harbinger of Alzheimer dementia Archives of Otolaryngology - Head and Neck Surgery 2011;137(4):390-5 doi: 10.1001/archoto.2011.2
105. Lister J, Roberts R, Shackelford J, Rogers C An adaptive clinical test of temporal resolution American Journal of Audiology 2006;15:133-40 doi: 10.1044/1059-0889(2006/017)
106. Lister J, Roberts RA, Krause JC, DeBiase D, Carlson H An adaptive clinical test of temporal resolution: Within-channel and across-channel gap detection International Journal of Audiology 2011;50:375-84 doi: 10.3109/14992027.2010.551217
107. Lister J, Roberts RA, Lister FL An adaptive clinical test of temporal resolution: Age effects International Journal of Audiology 2011;50(6):367-74 doi: 10.3109/14992027.2010.551218
108. Cox R A structured approach to hearing aid selection Ear and Hearing 1985;6(5):226-39
109. Delis DC, Kaplan E, Kramer JH. Delis-Kaplan Executive Function System (D-KEFS). San Antonio, TX: The Psychological Corporation; 2001.
110. Ruff RM, Light RH, Parker SB Benton Controlled Oral Word Association Test: Reliability and updated norms Archives of Clinical Neuropsychology 1996;11:329-38
111. Duff Canning SJ, Leach DL, D. S, Ngo L, Black SE Diagnostic utility of abbreviated fluency measured in Alzheimer disease and vascular dementia Neurology 2004;62:556-62 doi: 10.1212/WNL.62.4.556
112. Strauss ME, Sherman EM, Spreen O. A compendium of neuropsychological tests: Administration, norms, and commentary. 3rd ed: Oxford University Press; 2006.

113. Reitan RM, Wolfson D. The Halstead-Reitan Neuropsychological Test Battery. Tucson, AZ: Neuropsychological Press; 1985.
114. Lezak MD. Neuropsychological Assessment. 3rd ed. New York: Oxford University Press; 1995.
115. Owsley C, Sloane ME, McGwin G, Ball KK Timed instrumental activities of daily living tasks: Relationship to cognitive function and everyday performance assessments in older adults Gerontology 2002;48(4):254-65 doi: 10.1159/000058360
116. Robertson IH, Ward T, Ridgeway V, Nimmo-Smith I The test of everyday attention (TEA) Test Reviews 2001;4(1):51-5
117. Robertson IH, Ward T, Ridgeway V, Nimmo-Smith I. The Test of Everyday Attention Manual: Pearson Assessment; 1994. 32 p.
118. Gordon EE. Advanced Measures of Music Audiation. Chicago, IL: GIA Publications; 1989.
119. Sherer M, Maddux JE, Mercandante B, Prentice-Dunn S, Jacobs B, Robers RW The self-efficacy scale: Construction and validation Psychological Reports 1982;51:663-71
120. Zelenak M Development and validation of the Music Performance Self-Efficacy Scale Music Education Research International 2010;4:31-43
121. Bugos JA, Kochar S, Maxfield N Intense piano training on self-efficacy and physiological stress in aging Psychology of Music 2016:1-14 doi: 10.1177/0305735615577250
122. Rabipour S, Davidson PS Do you believe in brain training? A questionnaire about expectations of computerized cognitive training Behavioural Brain Research 2015 doi: 10.1016/j.bbr.2015.01.00225591472
123. Faul F, Erdfelder E, Lang AG, Buchner A G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences Behavior Research Methods 2007;39(2):175-912007/08/19 doi: 10.3758/BF0319314617695343
124. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.
125. Holm S A simple sequentially rejective multiple test procedure Scandinavian Journal of Statistics 1979;6:65-70
126. Roderick JA, Rubin DB, editors. Statistical Analysis with Missing Data. 2nd ed2002.
127. Hayes AF. PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]. 2012.
128. Selig JP, Preacher KJ Mediation models for longitudinal data in developmental research Reseach in Human Development 2009;6(2-3):144-64
129. Fritz MS, MacKinnon DP Required sample size to detect the mediated effect Psychological Science 2007;18:223-39 doi: 10.1111/j.1467-9280.2007.01882.x
130. Chermak GD, Musiek FE. Central auditory processing disorders: New perspectives. San Diego, CA: Singular1997.

13. RESULTS

Recruitment and enrollment began in January 2018 and continued through December 2020, with data collection completed in April 2021. Seventy-six percent of randomized participants completed post-test. Nineteen participants were excluded from analyses due to a head injury (*n*=1), undergoing anesthesia (*n*=10), or hospitalization (*n*=8) before post-test. Please see Figure 1 for details.

The randomized groups did not differ in sex, $\chi^2(1)=.787$, $p=.375$, ethnicity $\chi^2(1)=.167$, $p=.683$, or race, $\chi^2(1)=.045$, $p=.832$. The Music Reading Assessment scores did not differ between the two groups, $t(266)=-.958$, $p=.338$. The randomized groups did not significantly differ in age, $F(1,266)=.293$, $p=.589$, education, $F(1,266)=.369$, $p=.544$, MoCA, $F(1,266)=.427$, $p=.514$, or GDS, $F(1,266)=.943$, $p=.332$. However, there were marginally significant group differences for PTA in the left ear, $F(1,266)=3.31$, $p=.070$ and statistically significant group differences for both PTA in the right ear, $F(1,266)=6.69$, $p=.010$, and visual acuity, $F(1,266)=8.99$, $p=.003$. The music listening group tended to have worse hearing, while the piano training group tended to have worse visual acuity. See Table 1. The groups also differed significantly in their expectations regarding potential benefits of their randomized group, $F(1,214)=15.07$, $p<.001$. The piano training group had an average rating of 5.55 while the music listening group had an average rating of 4.99, indicating the piano training condition had higher expectations about the potential effects of their randomized condition. The piano training group also rated their condition as more challenging with 92% agreeing that the intervention was challenging as compared to 86.3% of the music listening group indicating their condition was challenging ($p<.001$).

Principal component analysis with varimax rotation was conducted to reduce the number of variables for analyses. Please see Supplemental Table A for details. The first factor reflected cognitive performance speed and included the Trail Making Test, Digit Coding, and the TEA visual elevator and phone search subtests. The second factor included all three verbal fluency subtests. The third factor reflected CAP and included DDT, DSI, TCS, and WIN. The fourth factor included two indices from the TEA phone search while counting and visual elevator accuracy. The fifth composite included ATTR across- and within- channel performance. Baseline composites of cognitive performance, verbal fluency, ATTR, CAP, and everyday function were created by averaging z scores after the reverse-scaling of items with negative factor loadings. Post-test scores were standardized by baseline means and *SD*s. Timed IADL performance did not significantly load on any factor and was thus examined separately.

In the primary analyses, we examined whether performance changed differentially by random assignment across composite outcomes from baseline to post-test as indicated by a significant group-by-time interaction. Covariates included hearing PTA, visual acuity, and expectations. The group-by-time interactions were not statistically significant for any of the outcomes (*p*s>.194). See Table 2. The results did not change when we included only those who were adherent (*p*s>.180). See Supplementary Table B. Piano training did not significantly enhance CAP, cognition, or everyday function as compared to music listening. We further examined if intervention effects varied by MCI status (Supplementary Table B). There were no significant group-by-time-by-MCI status interactions indicating those with and without MCI did not show differential benefit (*p*s>.245). Thus, MCI did not significantly moderate the effects of piano training.

Effect sizes for piano training as compared to music listening are shown in Table 3. Improvements were not evident in the primary analyses, adherent analyses, or the subsample without MCI on measures of CAP, cognition, or everyday function. Participants with MCI randomized to piano training showed potential small effect sizes for improvement relative to music listening on Trails A, Digit Symbol Coding, and one subtest of the TEA (*d*s>=0.25). Raincloud plots of individual assessment and composite variables are included in the supplemental materials.

Results indicated significant effects of time for ATTR in all models and significant effects of time for CAP in the primary and adherent analyses indicating improved performance from pre- to post- training. Significant effects of time were found for cognition and TEA in the primary and adherent analyses, indicating a tendency for decline across time. With respect to covariate effects, CAP, ATTR, and cognitive performance were significantly impacted by hearing. Similarly, Timed IADL and cognitive performance varied significantly by visual acuity. Those with poorer hearing or vision tended to perform worse on these outcomes. Interestingly, expectations about the assigned condition were significantly related to composite outcomes of CAP and cognition. However, having greater expectations that the randomized condition would positively affect abilities was

associated with poorer performance. See Table 2. Additional sensitivity analyses examined if there were differential effects by sex. No significant group-time-sex interactions were evident (*p*s>.119), and the pattern of results was the same when analyses were stratified by sex.

Table 1. *Participant characteristics by randomized arm and cognitive status*.

	Piano Training		Music Listening	
Cognitive status	Non-MCI	MCI	Non-MCI	MCI
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
	n (%)	n (%)	n (%)	n (%)
Age in years	69.1 (5.5)	71.4 (7.1)	69.6 (5.2)	72.1 (5.7)
Education in years	16.0 (2.3)	16.5 (2.2)	16.0 (2.1)	15.4 (2.3)
Sex (% female)	63 (65%)	18 (48%)	58 (59%)	17 (44%)
Music Reading Assessment score (/40)	2.52 (4.31)	2.43 (3.83)	3.42 (5.41)	2.08 (4.29)
Montreal Cognitive Assessment score (/30)	26.6 (2.1)	23.3 (1.4)	26.6 (1.9)	22.8 (1.6)
Geriatric Depression Scale short-form score (/15)	0.9 (1.2)	1.1 (1.1)	0.9 (1.1)	0.6 (0.9)
Pure tone average left ear (dB)+	19.8 (9.4)	20.4 (9.4)	20.9 (10.8)	27.7 (15.9)
Pure tone average right ear (dB)+*	18.6 (10.3)	20.4 (9.9)	22.4 (14.0)	25.8 (11.6)
Visual acuity (logMAR)+*	.04 (.12)	.09 (.15)	.01 (.13)	.01 (.10)
Expectations (rating/7)*	5.4 (1.2)	5.8 (0.8)	4.9 (0.8	5.3 (0.8)

Note. MCI=mild cognitive impairment. +Lower scores reflect better performance. *Significant differences by randomized group were evident in pure tone average right ear, visual acuity, and expectations (*p*s<.01).

Table 2. *The effects of randomization to piano training as compared to music listening from baseline to post-training across composite outcomes of central auditory processing, cognition, and everyday function.*

	Estimate	SE	*p*	Estimate	SE	*p*
	ATTR+			Central auditory processing		
Intercept	53.90	1.17	<.001	46.25	0.73	<.001
Group	-2.24	2.37	.346	-1.72	1.49	.249
Time	-2.77	0.68	<.001	2.29	0.36	<.001
Group x time	1.51	1.35	.265	.465	0.71	.515
Hearing	0.12	0.06	.030	-0.63	0.04	<.001
Visual acuity	6.24	4.71	.187	-7.270	3.97	0.07
Expectations	0.79	0.59	.185	-1.20	0.49	.016
	Verbal fluency			Cognition		
Intercept	49.32	0.98	<.001	51.81	0.90	<.001
Group	-1.81	2.00	.367	-2.62	1.84	.157
Time	0.41	0.48	.393	-1.40	0.42	.001
Group x time	-0.29	0.96	.763	0.39	0.84	.638
Hearing	-0.09	0.06	.134	-0.14	0.06	.016
Visual acuity	-6.55	5.33	.220	-15.45	5.12	.003
Expectations	-0.79	0.67	.234	-1.96	0.64	.003
	Timed IADL+			Test of Everyday Attention		
Intercept	48.89	1.32	<.001	62.34	1.15	<.001

Group	3.15	2.67	.240	1.37	2.33	.557
Time	0.14	0.79	.863	-8.61	0.67	<.001
Group x time	-1.53	1.58	.327	-0.40	1.340	.765
Hearing	0.07	0.57	.194	-0.02	0.05	.603
Visual acuity	12.75	4.770	.008	0.24	4.470	.957
Expectations	1.02	0.59	.089	-0.79	0.56	.157

Note. ATTR=Adaptive Tests of Temporal Resolution, IADL=instrumental activities of daily living. MCI=mild cognitive impairment. +For ATTR and Timed IADL lower scores reflect better performance.

Table 3. *Effect sizes for piano training relative to music listening across composite outcomes and assessments for the analytic sample overall, among those adherent, and by cognitive status*.

Outcome Measure	*d* Effect Size			
	Overall (*n*=198)	Adherent (*n*=189)	Non-MCI (*n*=146)	MCI (*n*=52)
Central auditory processing composite	-0.01	-0.01	0.00	-0.07
Dichotic Digits	-0.06	-0.07	-0.03	-0.14
Dichotic Sentence Identification	0.01	0.01	0.02	-0.07
Time Compressed Speech	-0.01	0.01	0.03	-0.06
Words in noise	0.03	0.03	0.12	0.01
Adaptive Tests of Temporal Resolution composite	0.17	0.20	-0.09	0.24
Cognitive composite	-0.03	-0.03	-0.10	0.08
Trails A	0.03	0.06	0.00	0.43
Trails B	-0.17	-0.20	0.01	-0.54
Digit Symbol Coding	0.08	0.09	0.06	0.26
Test of Everyday Attention - Visual elevator timing subtest	0.02	0.02	0.03	0.44
Test of Everyday Attention - Telephone search time	-0.29	-0.32	0.08	-0.27
Verbal fluency composite	0.05	0.06	0.09	-0.12

Letter fluency	0.00	0.02	0.09	-0.36
Category fluency	0.18	0.20	0.18	0.13
Switching fluency	-0.06	-0.05	-0.07	-0.05
Test of Everyday Attention composite	0.04	0.08	-0.30	-0.02
Visual Elevator accuracy	0.01	0.03	-0.12	-0.19
Dual task decrement	0.10	0.13	-0.18	0.22
Timed Instrumental Activities of Daily Living composite	-0.15	-0.21	0.09	-0.13

Note. MCI=mild cognitive impairment. *d* effect sizes were calculated as pre- to post- change in the piano training condition relative to the music listening condition divided by the baseline standard deviation such that positive values would indicate improvement. Adherent analyses include those who completed 15 or more sessions of the assigned condition.

Figure 1. CONSORT Flow Diagram