

Original Research Report

A Community Choir Intervention to Promote Well-Being Among Diverse Older Adults: Results From the Community of Voices Trial

Julene K. Johnson, PhD,^{1,2,*} Anita L. Stewart, PhD,^{1,2} Michael Acree, PhD,³ Anna M. Nápoles, PhD,⁴ Jason D. Flatt, PhD,¹ Wendy B. Max, PhD,¹ and Steven E. Gregorich, PhD^{2,5}

¹Institute for Health & Aging, ²Center for Aging in Diverse Communities, and ³Osher Center for Integrative Medicine, University of California San Francisco. ⁴Intramural Research Program, National Institute of Minority Health and Health Disparities, National Institutes of Health, Bethesda, Maryland. ⁵Division of General Internal Medicine, Department of Medicine, University of California San Francisco.

*Address correspondence to: Julene K. Johnson, PhD, Institute for Health & Aging, University of California San Francisco, 3333 California Street, San Francisco, CA 94118. E-mail: Julene.johnson@ucsf.edu

Received: June 14, 2018; Editorial Decision Date: October 29, 2018

Decision Editor: Derek Isaacowitz, PhD

Abstract

Objectives: To test effects of the Community of Voices choir intervention on the health, well-being, and health care costs of racial/ethnically diverse older adults.

Method: Twelve Administration-on-Aging-supported senior centers were cluster randomized into two groups: the intervention group started the choir immediately and a wait-list control group began the choir 6 months later. The choir program was designed for community-dwelling adults aged 60 years and older. The multimodal intervention comprises activities that engage participants cognitively, physically, and socially. Outcome measures assessed these three domains as well as health care utilization and costs. The intention-to-treat comparison was at 6 months.

Results: The sample ($N = 390$) had a mean age of 71.3 years ($SD = 7.2$); 65% were nonwhite. Six-month retention was 92%. Compared to controls, intervention group members experienced significantly greater improvements in loneliness ($p = .02$; standardized effect size [ES = 0.34] and interest in life ($p = .008$, ES = 0.39). No significant group differences were observed for cognitive or physical outcomes or for health care costs.

Discussion: Findings support adoption of community choirs for reducing loneliness and increasing interest in life among diverse older adults. Further efforts need to examine the mechanisms by which engagement in choirs improves aspects of well-being and reduces health disparities among older adults, including potential longer-term effects.

ClinicalTrials.gov Registration: NCT01869179 registered January 9, 2013.

Keywords: Creativity, Health promotion, Minority and diverse populations, Music

The United States' population of adults aged 65 years and older is increasingly diverse, with nearly 22% being from racial/ethnic minority backgrounds (U.S. Census Bureau, 2017); by 2030, this percent will increase to almost one

third (Vincent & Velkoff, 2010). In addition, many older adults face socioeconomic challenges, for example, more than 4.2 million were below the poverty level in 2015 (Administration on Aging, 2016). Among older adults, the

poverty rate is higher for blacks (18.4%), Latinos (17.5%), and Asians (11.8%) compared to non-Latino whites (6.6%). Minority and lower socioeconomic status (SES) older adults are at greater risk for poor health outcomes (Louie & Ward, 2011).

Engagement in the arts is one novel approach to improve the health of diverse older adults. Arts-based interventions can be offered in the community, are relatively low cost to deliver, are engaging, and can be culturally tailored (Macdonald, Kreutz, & Mitchell, 2012; Noice, Noice, & Kramer, 2014; Castora-Binkley, Noelker, Prohaska, & Satariano, 2010). Because music is integral to most cultures, music interventions may provide opportunities for diverse older adults to remain active and engaged (Johnson et al., 2015). Community choirs are popular, with approximately 32.5 million adults singing regularly in choirs in the United States. (Chorus America, 2009). Community choirs also have few requirements in terms of musical ability because basic singing abilities develop spontaneously during childhood (Trainor & Hannon, 2013). Choirs are multimodal, defined as having cognitive, physical, and psychosocial engagement components. Participating in activities that involve multiple engagement components in later life and that occur in groups may confer more health benefits than activities with only one component or done alone (Karp et al., 2006; Menec, 2003).

There is a need to understand the potential health benefits of participating in community choirs in later life. Several cross-sectional studies found that older adults who participate in choirs report greater well-being (Clift et al., 2010; Gick, 2011; Johnson, Louhivuori, & Siljander, 2017). However, these studies involved older adults who were predominantly white and from relatively high SES backgrounds. The first longitudinal study was nonrandomized and compared older adults who participated in a 12-month choir to a usual activity group (Cohen et al., 2006). Participants who sang in the choir reported better self-rated health, fewer doctor visits, use of fewer over-the-counter medications, fewer falls, and less decline in morale compared to the usual activity control group. However, the groups were self-selected, bias due to attrition was not considered, and the sample was predominantly white women. Recently, a randomized controlled trial of a choir program for community-dwelling older adults was conducted in England (Coulton, Clift, Skingley, & Rodriguez, 2015). The program involved singing 90 min a week for 14 weeks, and health and economic outcomes were evaluated at the end of the intervention. At 3 months, the choir group had significantly higher scores on a mental health-related quality of life measure and lower depression and anxiety compared with a usual activity control group. No differences were observed on a physical health-related quality of life measure. Participants were predominantly white women with high educational backgrounds. Although none of the studies reviewed earlier examined cognitive outcomes, studies of piano training found improvements in cognition for older adults (Bugos, Perlstein, McCrae, Brophy, & Bedenbaugh, 2007), and one

recent qualitative study reported improvements in cognitive function as a perceived benefit of choir singing (Fu, Lin, Belza, & Unite, 2015). Additional review of choir studies are in a prior article (Johnson et al., 2015). Thus, well-designed studies are needed to determine whether choir singing might have effects on multiple aspects of health and well-being for more diverse samples.

This article reports the effects of a community choir intervention designed to promote the health and well-being of diverse older adults. The study used community-engaged research methods (Napoles, Santoyo-Olsson, & Stewart, 2013) to design and implement a community choir intervention. To evaluate the choir intervention, we incorporated an array of outcomes measuring physical, cognitive, and psychosocial aspects of health. Here, we report the main randomized intention-to-treat group comparisons at 6 months. Compared to delayed intervention controls, we hypothesized that participation in a community choir program will be associated with improvements/maintenance on the primary outcomes of executive function, depressive symptoms, and lower body strength at 6 months. We also hypothesized that, compared to delayed intervention controls, participation in a community choir program will be associated with improvements/maintenance in the secondary outcomes of cognition (attention and inhibitory control, memory), psychosocial well-being (sadness, positive affect, fear/affect, loneliness, interest in life), physical (balance, gait speed), and health care utilization and costs at 6 months.

Methods

The Community of Voices (COV; Comunidad de Voces in Spanish) study was a multisite, cluster-randomized trial conducted at 12 Administration-on-Aging-supported senior centers serving racial/ethnically diverse communities throughout San Francisco, CA (study enrollment February 2012–August 2015). Details about the study design, senior centers, and recruitment methods are published elsewhere (Johnson et al., 2015, 2017). Twelve Administration-on-Aging-supported senior centers were randomized to receive the choir intervention immediately (intervention group) or after a 6-month delay (wait-list control group). Although each choir spanned 1 year (44 sessions), the main randomized intention-to-treat group comparison reported here was at 6 months (after 23 sessions) because this coincided with the start of the choir intervention for the control group.

Inclusion/Exclusion Criteria

Inclusion criteria were being aged 60 years and older, having sufficient visual and hearing acuity (with assistive devices), and being fluent in English or Spanish (bilingual and monolingual speakers). Inclusion criteria were intentionally broad because this was a community-based effectiveness trial. Exclusion criteria included having cognitive impairment (Mini-Cog score <3 of 5; Borson, Scanlan,

Brush, Vitaliano, & Dokmak, 2000), self-reported diagnosis of dementia, a serious medical condition, or plans to move out of the area. Persons who were regularly singing in a choir during the past 6 months (e.g., weekly) were excluded. The study was approved by the University of California, San Francisco's Institutional Review Board protocol number 12-09005, and written consent was obtained.

Randomization Procedures

We used a two-arm cluster randomized trial design with the 12 senior centers serving as clusters. A restricted randomization procedure (Hayes & Moulton, 2009) allocated the 12 centers into 2 sets of 6 centers to optimize cross-set balance of 78 center characteristics: 10 indicators of services/activities provided (e.g., meal service, transportation), 16 demographic indicators (e.g., proportional representation of age strata, racial/ethnic groups, sex, functionally impaired), and center readiness to implement the intervention. Afterward, one set of six centers was randomly assigned to the intervention group (began choir immediately) and the other set to the wait-list control group (waited 6 months to begin the choir). Recruitment occurred in phases, one pair of senior centers at a time, where each pair consisted of one intervention and one wait-list control center. Once baseline assessments were completed at a pair of sites, randomization assignment was revealed to study and senior center staff, music professionals, and participants.

Intervention: COV Choir Program

The choir program was designed in collaboration with a community music partner and the senior centers with the goal of promoting the health and well-being of diverse, community-dwelling older adults. The theoretical framework and rationale for the program content are published elsewhere (Johnson et al., 2015). Each choir session included activities targeting three hypothesized pathways by which a choir could promote health and well-being: cognitive, physical, and psychosocial engagement. The approach to the choir intervention across sites was standardized around these components and documented in a manual (<https://cov.ucsf.edu/>). The choirs were led by professional choir directors and accompanists from local communities who completed training on the intervention components prior to starting the choir with one refresher training midway. Choir directors identified music repertoire that could be culturally tailored for each site and was appropriate for older adults with a range of singing abilities and experience and challenging enough to facilitate growth and mastery over time. The 90-min choir sessions took place at the senior centers, and each choir met weekly for 44 weeks, including 3–4 informal public performances. Additional details about the intervention are found in the online manual.

Fidelity for Delivery of Intervention

The music partner's program manager helped supervise delivery of the choir intervention. The principal investigator completed fidelity checks using a 23-item survey that assessed three areas: implementation of the three choir components (i.e., cognitive, physical, and psychosocial engagement), leadership skills, and musicianship. Each item was rated on a 4-point scale assessing the extent to which the director met expectations (3 exceeded expectations, 2 met expectations, 1 below expectations, and 0 well below expectations). Fidelity checks were completed 3 weeks after the choir started, at 3 months, and at 6 months. If ratings on any item were below expectations (0 or 1), feedback was provided and a follow-up fidelity visit was completed within 2 weeks. We report the percentage of items that were rated as having met expectations (a rating of 2 or 3) for each of the fidelity areas at all three time points.

Participant Adherence to Intervention

To assess participant adherence, attendance at each choir session was recorded by the choir directors. Attendance was calculated as the total number (and percent) of the 23 sessions attended and percent of participants who completed at least 50% and at least 75% of the sessions.

Assessments

In-person assessments were conducted at each of the senior centers. We used the National Institutes of Health (NIH) Assessment Center computer-based platform to collect and manage all data and added several project-specific instruments.

Sample characteristics

Demographic variables included age (years), sex (male/female), ethnicity (Hispanic/Latino or not), race (American Indian/Alaskan Native, Asian, black/African American, Native Hawaiian/other Pacific Islander, white, and other), educational level, marital status, language of screening interview (Spanish/English), and nativity. We assessed financial hardship in terms of difficulty paying for food, monthly bills, medical visits, or prescribed medications in the past 12 months. We assessed health insurance (none, public only, or any private), whether they previously sang in a choir as an adult, and their overall music ability (poor, fair, good, very good, or excellent).

Self-rated health was assessed using a standard item with five response choices (excellent, very good, good, fair, poor). We asked whether a health professional had told them they had any of 11 chronic health conditions, which were categorized into eight physical conditions (diabetes, hypertension, stroke, cancer, arthritis, and emphysema/bronchitis/asthma) and one mental health condition (depression or anxiety).

Outcome measures

As detailed elsewhere (Johnson et al., 2015), outcome measures were selected based on the hypothesized mechanisms of action of the intervention by engagement domain: psychosocial, cognitive, and physical. We included measures from the NIH Toolbox for the Assessment of Neurological and Behavioral Function (NIH Toolbox; Hodes, Insel, & Landis, 2013) as well as several commonly used legacy measures. For each engagement domain, we selected one measure a priori as the primary outcome; other measures of that category were secondary outcomes. Here, we briefly describe each measure in terms of its content; details about the response choices, scoring, and reliability are reported elsewhere (Johnson et al., 2015, 2017).

Psychosocial outcomes

We assessed six psychosocial outcomes: depressive symptoms, sadness, anxiety, loneliness, positive affect, and interest in daily life. The eight-item Patient Health Questionnaire (PHQ-8; Kroenke et al., 2009) assessing depressive symptoms was the primary psychosocial outcome and ranges from 0 to 24 with higher scores indicating more depression. The other five measures were drawn from the NIH Toolbox. Three assessed psychological distress: Sadness, Fear/Affect (symptoms of anxiety), and Loneliness. Two measures assessed psychological well-being: Positive Affect and Apathy (Interest in Life). We reversed and renamed "Apathy" as "Interest in Life" because all items pertain to interest in life (e.g., interested in things, got things done, did interesting things, and was motivated). For NIH Toolbox measures, we used T-scores scaled so that the means equaled 50 and the standard deviations equaled 10.

Cognitive outcomes

Three tests assessed memory and executive function. The Trail Making Test (TMT) was the primary cognitive outcome. The TMT (Reitan, 1958; Strauss, Sherman, & Spreen, 2006) parts A and B were used to assess the set-shifting component of executive function. We allowed a maximum time of 180 s to complete each part to reduce participant burden (Tombaugh, 2004). The variable of interest is time (seconds) to complete TMT-B minus time to complete TMT-A, which is an index of executive function that isolates the executive control component (Sanchez-Cubillo et al., 2009). The NIH Toolbox Flanker Inhibitory Control and Attention Test was used as a test of attention and inhibitory control. Participants completed 20 trials; the final computed score reflects a combination of accuracy and reaction time. A modified version of the NIH Toolbox Rey Auditory Verbal Learning Test was used as a measure of episodic memory. Fifteen unrelated words were presented orally (via audio recording) over three consecutive trials. After each presentation, participants were asked to recall as many words as possible and again after a distractor trial, which included 15 new words. Immediately after recall of the distractor list, participants were asked to recall as many

of the original 15 words as possible (delayed recall, possible score 0–15), which was the variable of interest.

Physical outcomes

Three performance-based measures were used to assess lower body strength, balance, and walking speed. Chair stands from the short physical performance battery (Guralnik et al., 1994) assessed the primary physical outcome of lower body strength. We recorded the time in seconds to complete five chair stands, allowing a maximum of 60 s. The NIH Toolbox Standing Balance measure (Reuben et al., 2013) assessed static standing balance. Participants were asked to maintain five poses ranging from standing with feet side by side to a tandem pose. Originally we planned to use the ratio of poses 4:1, but laptop connectivity problems (i.e., transmission errors between the accelerometer and the computer accessing the assessment center) resulted in a large amount of missing data. Here, we report results for the ratio of poses 2:1. The NIH Toolbox performance measure of gait speed (Reuben et al., 2013), which requires participants to walk 4 m at their usual pace, was used as a measure of walking pace. The faster of two trials was the variable of interest.

Health care utilization and costs

At baseline and 6 months, we collected self-reported health care utilization over the prior 3 months, in person or by phone using a modified measure from the Chronic Disease Self-Management study (Lorig et al., 1996). We assessed number of visits to a doctor, mental health provider (e.g., counselor, psychologist), and other health providers (e.g., home health nurse, physical therapist); number of emergency room visits and hospitalizations (number of nights hospitalized); and number of outpatient surgeries over the past 3 months. The mean cost per unit of utilization (e.g., doctor visit, hospital night) was obtained from an analysis of the 2014 Medical Expenditures Panel Survey for adults aged 65 years and older and multiplied by the number of units used to obtain health care costs. By converting health care utilization to costs, more weight is given to more costly services (e.g., hospitalizations and outpatient surgeries) and the substitution of cheaper services for more expensive services is taken into account.

Power Analysis

Original assumptions included intention-to-treat analyses, 80% power, two-tailed alpha of 0.05, $N = 450$ enrolled patients across 12 senior centers, 80% retention at 6 months ($n = 360$), casewise deletion of missing data (for power analysis, only), and linear mixed models of group differences at 6 months. On the basis of our preliminary data on choral groups, the largest intracluster correlation (ICC) for any outcome equaled 0.015. Accounting for ICC, cluster size, and attrition, the minimum detectable 6-month standardized group difference, d , equaled 0.35, which

compared favorably to effect sizes based upon the results reported by [Cohen and colleagues \(2006\)](#) for overall health rating (0.41) and depression (0.38) outcomes. We did not make alpha adjustments for testing the set of clinically distinct outcomes pertinent to the hypothesized mechanisms of this multimodal experimental intervention. Such adjustments presume a universal null hypothesis that holds for all outcomes simultaneously but, because we cannot prespecify which outcome or outcomes may most influence subsequent COV-related policy decisions, the universal null is not of primary interest ([Cook & Farewell, 1996](#); [Cox, 1965](#); [Perneger, 1998](#); [Rothman, 1990](#)). Instead, we specified a test-wise error rate to allow marginal inferences. This, in combination with the reported effect size estimates, should allow readers to draw conclusions about the impacts of the COV intervention on the modeled outcomes.

Data Analyses

The longitudinal data had a three-level nested structure: senior centers, participants, and repeated assessments. Linear mixed models with random intercepts of the cognitive, psychosocial, and most physical functioning outcomes were fit to the baseline and 6-month outcome assessments and tested intention-to-treat group, time, and group-by-time effects. The group-by-time interaction was the primary intervention effect test. Model-predicted means are reported for the cognitive, psychosocial, and most physical functioning outcomes. Total costs were modeled using a generalized estimating equation negative binomial model testing group-by-time interaction effects as the primary intervention effect test.

Results

Recruitment and Retention

During the study design phase, 12 senior center clusters were identified; all 12 clusters were eligible, agreed to participant, were randomized, and completed the study. Study recruitment took place over 29 months, and details are reported elsewhere ([Johnson et al., 2017](#)). [Figure 1](#) shows the Consolidated Standards of Reporting Trials diagram for participant flow. Briefly, six senior centers were randomized to the choir intervention group and six to the wait-list control group. Within those 12 centers, 636 individuals began the screening process, and 390 (61%) enrolled in the study ($n = 208$ in the choir intervention group and $n = 182$ in the wait-list control group). The study retention rate at 6 months was 90% for the intervention group and 93% for the control group (91.5% overall).

Sample Characteristics

Baseline demographic characteristics are summarized by group in [Table 1](#) and in [Supplementary Tables 1 and 2](#).

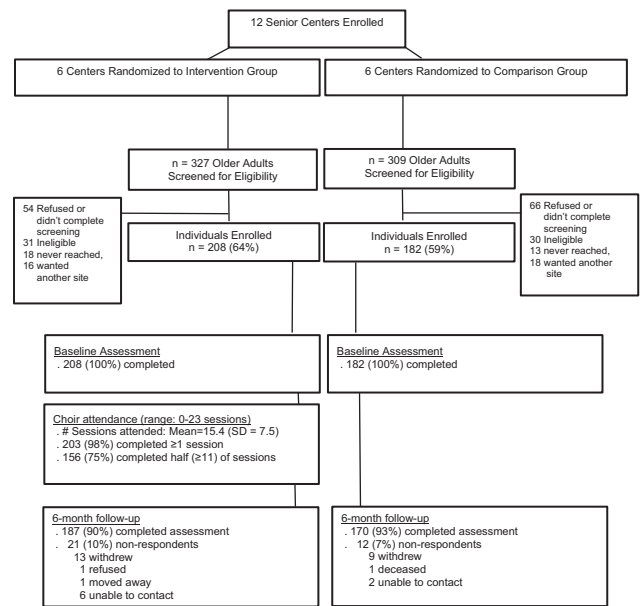


Figure 1. Consolidated Standards of Reporting Trials (CONSORT) diagram.

Overall, the mean age was 71.3 ($SD = 7.2$) years; 76% were female, and 65% reported being from nonwhite (minority) racial/ethnic backgrounds. Forty-one percent of the sample was born outside the United States. Approximately one third was married or living with a partner. A range of education was represented. Financial hardship was reported by 20%. Approximately 25% reported fair or poor health, and 60% reported two or more chronic medical conditions. More than half (55%) reported that they had not previously sung in a choir as an adult, and over half (56%) rated their musical ability as poor or fair. Demographic indicators were generally well balanced between groups.

Adherence to Intervention and Fidelity

During the first 6 months, participants in the choir intervention group attended an average of 15.4 ($SD = 7.5$) and a median of 19 of 23 sessions. Seventy-three percent of participants attended half or more of the sessions, and 58% attended three quarters or more of the sessions.

Fidelity checks found that the interventionists (directors) met expectations for 99% of musicianship items, 98% of leadership/communication items, and 83% of intervention program components. Of these, ratings on the physical engagement domain were lower than for psychosocial and cognitive components.

Intention-to-Treat Results

[Table 2](#) summarizes the intention-to-treat results for baseline and 6-month outcomes. There were no significant group-by-time differences at 6 months on the three primary outcome measures: PHQ-8, TMT, or chair stands

Table 1. Sample Characteristics

Characteristics, N (%) ^a or mean (SD)	Control	Intervention
Age (range = 59–93) ^b , years	70.5 (7.3)	71.8 (6.9)
Sex (female)	140 (77)	158 (76)
Race/ethnicity		
Non-Latino white	51 (28)	85 (41)
Non-Latino black	62 (34)	42 (20)
Asian	45 (25)	33 (16)
Latino	24 (13)	48 (23)
High school or less education	49 (27)	54 (26)
Married or partnered	58 (32)	67 (32)
Financial hardship ^c	36 (20)	41 (20)
Foreign born	77 (42)	83 (40)
Mini-Cog total score (range 0–5) ^d	4.3 (1.0)	4.4 (0.9)
Self-reported fair/poor health	51 (28)	52 (25)
Diabetes	37 (21)	39 (20)
Hypertension	105 (59)	99 (52)
Heart disease	33 (19)	43 (23)
Stroke	15 (9)	12 (6)
Cancer	25 (14)	35 (18)
Arthritis	84 (48)	102 (53)
Pulmonary ^e	29 (16)	36 (19)
Depression or anxiety	37 (21)	52 (27)

Note: Control $n = 182$; Intervention $n = 208$; SD = standard deviation.

^aPercent of nonmissing.

^bOne participant aged 59 years was included in the study because of incorrect report of age at screening.

^cFinancial hardship = any “yes” to problems paying for food, bills, medical visits, or prescriptions.

^dTwo participants with a Mini-Cog score of 0 were included in the study per PI decision.

^eIncludes emphysema, bronchitis, and asthma.

($ps > .05$). However, significant group differences were found for two of six psychosocial outcomes. There were significant group-by-time interaction effects for Loneliness ($p = .02$; ES = 0.34 and Interest in Life ($p = .008$, ES = 0.39). There was a decrease in Loneliness in the intervention group from baseline to 6 months, whereas the control group did not differentially change. Interest in Life increased in the intervention group but not the control group. No significant group differences were observed for cognitive or physical outcomes ($ps > .05$).

Supplementary Tables 1 and 2 show the health care utilization and cost measures for the two groups at baseline and 6 months. Although total costs were similar for the two groups at baseline (\$1,317 for the intervention group vs \$1,300 for the control group), the costs for most services increased for both groups between baseline and 6 months. Whereas health care costs for the control group nearly tripled and costs for the intervention group doubled over the 6 month period, the group-by-time interaction effect was not statistically significant ($p = .440$).

Discussion

The COV trial represents the largest prospective randomized trial to date, testing the effects of a community

choir intervention on multiple domains of health and well-being of diverse older adults. The choir intervention improved two important factors affecting older adults' well-being: It reduced subjective feelings of loneliness and increased interest in life. However, cognitive and physical outcomes and three other psychosocial outcomes (depressive symptoms, anxiety, and positive affect) did not differentially change. Although the control group had nearly a 50% larger increase in health care costs than the intervention group, the difference was not statistically significant. A hallmark of the study was our ability to enroll and retain a large sample of diverse racial/ethnic and low-SES older adults who are typically underrepresented in health research. Another defining feature of the study was the use of community-engaged research methods whereby community partners were involved in intervention development and delivery.

The fact that participation in a choir reduced feelings of loneliness addresses a serious condition for older adults. Loneliness, defined as subjective and distressing feelings of isolation, less than desired amounts of companionship, and not belonging (Peplau, 1985) is common in older adults (Mezuk et al., 2016). It is well documented that older adults who experience feelings of loneliness are at increased risk for poor psychological well-being, cognitive decline, functional and motor decline, and death (Perissinotto, Cener, & Covinsky, 2012). In our conceptual framework (Johnson et al., 2015), we hypothesized that participation in a choir could help reduce loneliness by providing a meaningful and regular opportunity to meet new people, build social support, and increase one's sense of belonging. Our findings are consistent with recent reviews suggesting that opportunities for building social networks and/or support in a group format, including the arts, could help reduce loneliness among older adults (Cohen-Mansfield & Perach, 2015). What remains unclear is whether group singing facilitates social bonding above and beyond that conferred by other group activities (e.g., yoga, dancing; Pearce, Launay, Machin, & Dunbar, 2016). Future research should examine bonding in group singing and its effects on feelings of loneliness and isolation in older adults, particularly those from diverse backgrounds who may have different social networks and support systems (Smyth, Siriwardhana, Hotopf, & Hatch, 2015).

The COV intervention also increased interest in life (or reduced apathy). In healthy, nonclinical populations, apathy has been conceptualized as a reduction in goal-directed behaviors and loss of motivation (Bonnelle, Manohar, Behrens, & Husain, 2016). This is relatively common among community-dwelling older adults (Brodaty, Altendorf, Withall, & Sachdev, 2010) and is associated with incident cardiovascular disease, executive dysfunction, and changes in frontal lobe brain connectivity (Eurelings et al., 2014; Kawagoe, Onoda, & Yamaguchi, 2017). In our conceptual framework, we hypothesized that singing in a choir could increase interest in life by having a regular activity (Johnson et al., 2015). The fact that the choir met weekly provided the structure and opportunity

Table 2. Effect of Intervention and Control Groups on 6-month Changes From Baseline: Community of Voices Trial, San Francisco, CA

	High score ^a	Intervention		Control		Group × Time interaction		Standard effect size ^c
		Baseline	6 months	Baseline	6 months	Coefficient ^b	<i>p</i>	
Cognitive								
TMT ^d	-	73.7 (5.6)	74.8 (5.7)	79.7 (5.7)	81.0 (5.8)	-0.14	.97	0.01
Flanker ^e	+	7.3 (0.2)	7.2 (0.2)	7.1 (0.2)	7.1 (0.2)	-0.05	.65	0.07
RAVLT ^f	+	5.7 (0.3)	6.6 (0.3)	5.5 (0.3)	6.2 (0.3)	0.20	.41	0.12
Psychosocial								
PHQ-8	-	4.3 (0.3)	4.0 (0.3)	4.3 (0.3)	4.2 (0.3)	-0.19	.62	0.072
Sadness ^g	-	48.2 (0.6)	47.3 (0.6)	47.3 (0.6)	46.8 (0.6)	-0.42	.60	0.078
Positive affect ^h	+	48.8 (0.7)	49.4 (0.7)	51.3 (0.8)	50.5 (0.8)	1.41	.11	0.23
Fear/affect ^g	-	50.3 (0.7)	49.8 (0.7)	49.2 (0.7)	49.2 (0.7)	-0.49	.56	0.09
Loneliness ^h	-	52.1 (0.7)	50.4 (0.7)	49.6 (0.8)	50.1 (0.8)	-2.11	.02	0.34
Interest in life ⁱ	+	72.6 (0.4)	73.4 (0.5)	73.1 (0.5)	72.3 (0.5)	1.53	.008	0.39
Physical								
Chair stands ^j	-	12.1 (0.5)	12.5 (0.5)	13.4 (0.6)	13.1 (0.6)	-0.08	.86	-0.03
Balance ^k	-	1.5 (0.05)	1.4 (0.05)	1.5 (0.05)	1.4 (0.05)	-0.05	.52	0.10
Gait speed ^l	+	0.9 (.02)	0.9 (0.02)	0.9 (0.02)	0.9 (0.03)	0.00	.98	0.00

Note: Bold *p*-values indicate significant group differences. Bold outcome measures indicate primary outcomes; nonbold outcomes are secondary outcomes. PHQ-8 = eight-item Patient Health Questionnaire; RAVLT = Rey Auditory Verbal Learning Test; TMT = Trail Making Test.

^aDirection of high score: + = better, - = worse.

^bGroup × Time Interaction coefficient.

^cStandardized effect size, d_w , of the interaction effect. Here, d_w represents the group difference in within-person change expressed in units of the pooled within-person standard deviation (Hedges, 2007; Eq. 7).

^dTrail making-B minus Trail making-A (seconds).

^eComputed score.

^fRAVLT delayed recall.

^gT score

^hT score, Mean difference between intervention and control group at baseline, $p < .05$

ⁱAlthough labeled “apathy,” items pertain to interest in life (see measures description), T score.

^jSeconds to complete five chair stands.

^kStanding balance, ratio of position 2:position 1.

^l4-m walk, m/s.

for a regular activity. Several studies have found that participation in regular, structured activities as an older adult is associated with a reduced risk for cognitive decline, dementia, and functional decline (Stern & Munn, 2010; Verghese et al., 2003), but only a few studies involve diverse older adults (Wilson et al., 2003). Several studies suggest that older Latinos and blacks engage less frequently in activities than non-Latino whites (Herrera et al., 2011; Marquine, Segawa, Wilson, Bennett, & Barnes, 2012). In low-resourced communities, access to regular activities is important for older adults to remain active and engaged in their community, and participation in a choir could offer such opportunities.

We did not find significant group differences on physical or cognitive outcomes and discuss three possible explanations. First, it is possible that the physical and cognitive outcome measures selected for the study were not sensitive enough to detect change within the study follow-up interval or were not measuring the types of physical or cognitive

changes that may have occurred as a result of the intervention. The outcome measures in the current trial were modeled after other behavioral trials for older adults, including choirs. Second, participating in weekly 90-min sessions may not be a strong enough “dose” to affect cognitive and physical function of older adults, or it may take longer time periods for choir participation to affect physical and cognitive domains. Although we designed the COV intervention to be more engaging than typical choirs for older adults (e.g., sing-alongs) by requiring activities that involved increasingly challenging engagement, it is possible that the physical and cognitive engagement was not intense enough to produce changes in these domains. Perhaps offering a choir more than once a week along with daily singing tasks would be sufficient to achieve cognitive and physical maintenance or improvement. Further, the prevention of cognitive and physical decline has been difficult to document, for example, such as in studies of late-life engagement in other activities (e.g., cognitive training, physical activity, piano

playing; de Souto Barreto, Demougeot, Vellas, & Rolland, 2018; Butler et al., 2018; Schneider, Hunter, & Bardach, 2018). A third possible explanation is that the physical engagement activities, which were less familiar to music professionals, were the most inconsistently applied. Although choir directors completed training on each of the choir intervention components, more intensive training may be needed on physical engagement components.

Other multimodal interventions have also been evaluated. The Experience Corps intervention was similar to our choir intervention in that it aimed to simultaneously increase the social, physical, and cognitive engagement of the diverse older adults who volunteered in schools (Fried et al., 2013). Compared to low-intensity volunteering, older adults who were randomized to the intensive intergenerational volunteering intervention over two academic years experienced short-term improvements in executive function, increased walking activity, and enhancements in perceptions of generative desire and achievement (Carlson et al., 2009; Gruenewald et al., 2016; Varma et al., 2016). Importantly, the intervention required 15 hr/week of involvement over 2 years, which was more intense than our choir intervention. Interventions combining components such as physical activity, diet, and cognitive training have also been examined, but a recent report concluded that the evidence was insufficient to conclude that they benefit cognition (Kane et al., 2017). However, a large, long-term randomized multimodal intervention of diet, exercise, cognitive training, and vascular risk monitoring documented cognitive improvements after 2 years in at-risk older adults (Ngandu et al., 2015). It is possible that interventions aimed at improving cognition require sustained training periods.

Regarding costs, similar to the study by Coulton and colleagues (2015), we observed increased health care costs over time in both the choir intervention and control groups. It is difficult to know if rising costs reflect appropriate utilization (e.g., increased use and more intensive care because of chronic disease progression) or more care regardless of the need for care. Coulton and colleagues combined health and social care costs and did not find group differences in costs 3 months after their 14-week choir program ended.

There are some limitations of our study. Although the randomization assignment was kept confidential until all baseline assessments were completed at each pair of sites, intervention group assignment was known once the intervention began. Similar to most other choir studies, our study had a predominance of women. Further, our study was one of the early adopters of the NIH Toolbox in both Spanish and English, and we encountered some technical challenges in collecting data at senior centers. Nonetheless, its use resulted in more efficient administration because of computer-assisted testing and item response theory methods and reduced time for data entry.

In addition to being one of the first arts-based randomized trials for older adults, our pragmatic trial represents a new direction in translational research designed

to address health disparities in which interventions are designed and evaluated in community settings from the outset (Nápoles et al., 2013). First, the intervention was delivered in the community (at senior centers) by community music professionals. This approach optimizes the cultural relevance of the choir at each site and increases the likelihood of sustainability given that the choir was embedded in the community from the start. Second, study recruitment was conducted entirely in the community at senior centers, an approach that facilitated recruitment of an underrepresented population (Johnson et al., 2017), similar to other studies (Nápoles & Chadiha, 2011; Santoyo-Olsson et al., 2011). Third, all study assessments were conducted at the senior centers, making participation in the research convenient for participants and likely contributing to the excellent retention rate. These study methods can be a model for future trials to engage and retain diverse older adults in research. Conducting this trial in existing community settings moves the program further along on the translation continuum.

Our findings that singing in a choir can reduce feelings of loneliness and increase interest in life among diverse older adults have important public health implications. Community choirs are typically affordable, sustainable, and accessible, and can be culturally tailored, making them relevant and useful for helping to reduce health disparities among diverse older adults who are more likely to experience financial hardship and live in low-resourced communities compared with white older adults. Future research can focus on identifying novel mechanisms and measures to assess more nuanced impacts of community choirs on older adults' well-being.

Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* online.

Funding

This work was supported by the National Institute of Aging at the National Institutes of Health (R01AG042526 to J. K. Johnson and P30AG15272 to A. M. Nápoles/Karliner); and the National Center for Advancing Translational Sciences at the National Institutes of Health (UL1 TR000004). Dr. A. M. Nápoles' time was supported in part by the Intramural Research Program of the National Institute on Minority Health and Health Disparities. The contents and views in this article are those of the authors and should not be construed to represent the views of the National Institutes of Health, the funding sources, or the organizations with which they are affiliated.

Acknowledgments

We would like to acknowledge all of the research staff who helped with recruitment and collecting data: Jasmine Santoyo-Olsson, Rachel Freyre, Dana Pounds, Merima Ribic, Maria Cora, Ofelia

Villero, and Ariana Paniagua. We also thank the San Francisco Community Music Center choir directors and accompanists who delivered the Community of Voices choir intervention (Martha Rodriguez-Salazar, Jennifer Peringer, Maestro Curtis, Nola Curtis, Beth Wilmurt, Richard Daquioag, Billy Philadelphia, Helen Dilworth, Judy Lee, Leon Palad, and Eduardo Corzo), program manager Sylvia Sherman, and staff (Rachel Carlin, Christopher Borg, Sonia Caltvedt, Chus Alonso, Fran Hildebrand, and Stephen Shapiro). We would also like to acknowledge the San Francisco Department of Aging and Adult Services (Anne Hinton and Shireen McSpadden) and the participating senior centers: Mission Neighborhood Centers (Maria Bermudez), Centro Latino de San Francisco (Gloria Bonilla), Dr. George W. Davis Senior Center (Cathy Davis), Bayview Opera House (Barbara Ockel), Western Addition Senior Center (Robin Bill), San Francisco Senior Center-Aquatic Park (Sue Horst), 30th Street Senior Center (Valorie Villela), OMI Senior Center (Patty Clement-Cihak), Golden Gate Senior Services-Richmond Senior Center (Linda Murley), Golden Gate Senior Services-Castro Senior Center (Patrick Larkin), Bernal Heights neighborhood Center (Gina Dacus), IT Bookman Community Center (Jackie Wright and Kristin Rosboro), and Veterans Equity Center (Luisa Antonio). The study was made possible by these community partnerships.

Conflict of Interest

None reported.

Author Contributions

J. K. Johnson planned the study, supervised data collection and analysis, and helped write the article. A. L. Stewart helped plan the study, supervise data analysis, and write the article. M. Acree performed statistical analyses. A. M. Nápoles helped supervise data collection and analysis and write the article. W. B. Max helped plan the study, supervise data analysis, and write the article. J. D. Flatt helped with data analysis and writing. S. E. Gregorich helped plan the study, supervise data analysis, and write the article.

References

- Administration on Aging. (2016). *A profile of older Americans: 2016*. Washington, DC: U.S. Department of Health and Human Services.
- Bonnelle, V., Manohar, S., Behrens, T., & Husain, M. (2016). Individual differences in premotor brain systems underlie behavioral apathy. *Cerebral Cortex*, *26*, 807–819. doi:10.1093/cercor/bhv247
- Borson, S., Scanlan, J., Brush, M., Vitaliano, P., & Dokmak, A. (2000). The mini-cog: A cognitive “vital signs” measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry*, *15*, 1021–1027. doi:10.1002/1099-1166(200011)15%3A11<1021%3A%3A%3AID-GPS234>3.0.CO%3B2-6
- Brodaty, H., Altendorf, A., Withall, A., & Sachdev, P. (2010). Do people become more apathetic as they grow older? A longitudinal study in healthy individuals. *International Psychogeriatrics*, *22*, 426–436. doi:10.1017/S1041610209991335
- Bugos, J. A., Perlstein, W. M., McCrae, C. S., Brophy, T. S., & Bedenbaugh, P. H. (2007). Individualized piano instruction enhances executive functioning and working memory in older adults. *Aging & Mental Health*, *11*, 464–471. doi:10.1080/13607860601086504
- Butler, M., McCreedy, E., Nelson, V. A., Desai, P., Ratner, E., Fink, H. A., ... Kane, R. L. (2018). Does cognitive training prevent cognitive decline?: A systematic review. *Annals of Internal Medicine*, *168*, 63–68. doi:10.7326/M17-1531
- Carlson, M. C., Erickson, K. I., Kramer, A. F., Voss, M. W., Bolea, N., Mielke, M., ... Fried, L. P. (2009). Evidence for neurocognitive plasticity in at-risk older adults: The experience corps program. *Journals of Gerontology, Series A: Biological Sciences and Medical Sciences*, *64*, 1275–1282. doi:10.1093/gerona/glp117
- Castora-Binkley, M., Noelker, L., Prohaska, T., & Satariano, W. (2010). Impact of arts participation on health outcomes for older adults. *Journal of Aging, Humanities, and the Arts*, *4*, 352–367. doi:10.1080/19325614.2010.533396
- Chorus America. (2009). *The chorus impact study: How children, adults, and communities benefit from choruses*. Washington, DC: Chorus America.
- Clift, S. M., Hancox, G., Morrison, I., Hess, B., Kreutz, G., & Stewart, D. (2010). Choral singing and psychological well-being: Quantitative and qualitative findings from English choirs in a cross-national survey. *Journal of Applied Arts and Health*, *1*, 19–34. doi:10.1386/jaah.1.1.19/1
- Cohen, G. D., Perlstein, S., Chapline, J., Kelly, J., Firth, K. M., & Simmens, S. (2006). The impact of professionally conducted cultural programs on the physical health, mental health, and social functioning of older adults. *Gerontologist*, *46*, 726–734. doi:10.1093/geront/46.6.726
- Cohen-Mansfield, J., & Perach, R. (2015). Interventions for alleviating loneliness among older persons: A critical review. *American Journal of Health Promotion*, *29*, e109–e125. doi:10.4278/ajhp.130418-LIT-182
- Cook, R. J., & Farewell, V. T. (1996). Multiplicity considerations in the design and analysis of clinical trials. *Journal of the Royal Statistical Society Series A (Statistics in Society)*, *159*, 93–110. doi:10.2307/2983471
- Coulton, S., Clift, S., Skingley, A., & Rodriguez, J. (2015). Effectiveness and cost-effectiveness of community singing on mental health-related quality of life of older people: Randomised controlled trial. *British Journal of Psychiatry*, *207*, 250–255. doi:10.1192/bjp.bp.113.129908
- Cox, D. R. (1965). A remark on multiple comparison methods. *Technometrics*, *7*, 223–224. doi:10.2307/1266671
- de Souto Barreto, P. S., Demougeot, L., Vellas, B., & Rolland, Y. (2018). Exercise training for preventing dementia, mild cognitive impairment, and clinically meaningful cognitive decline: A systematic review and meta-analysis. *Journals of Gerontology, Series A: Biological Sciences and Medical Sciences*, *73*, 1504–1511. doi:10.1093/gerona/glx234
- Eurelings, L. S., Ligthart, S. A., van Dalen, J. W., Moll van Charante, E. P., van Gool, W. A., & Richard, E. (2014). Apathy is an independent risk factor for incident cardiovascular disease in the older individual: A population-based cohort study. *International Journal of Geriatric Psychiatry*, *29*, 454–463. doi:10.1002/gps.4026
- Fried, L. P., Carlson, M. C., McGill, S., Seeman, T., Xue, Q. L., Frick, K., ... Rebok, G. W. (2013). Experience corps: A dual trial to promote the health of older adults and children’s academic

- success. *Contemporary Clinical Trials*, *36*, 1–13. doi:10.1016/j.cct.2013.05.003
- Fu, M. C. C., Lin, S. Y., Belza, B., & Unite, M. (2015). Insights of senior living residents and staff on group-singing. *Activities, Adaptation & Aging*, *39*, 243–261. doi:10.1080/01924788.2015.1063332
- Gick, M. L. (2011). Singing, health and well-being: A health psychologist's review. *Psychomusicology*, *21*, 176–207. doi:10.1037/h0094011
- Gruenewald, T. L., Tanner, E. K., Fried, L. P., Carlson, M. C., Xue, Q. L., Parisi, J. M., ... Seeman, T. E. (2016). The Baltimore experience corps trial: Enhancing generativity via intergenerational activity engagement in later life. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *71*, 661–670. doi:10.1093/geronb/gbv005
- Guralnik, J. M., Simonsick, E. M., Ferrucci, L., Glynn, R. J., Berkman, L. F., Blazer, D. G., ... Wallace, R. B. (1994). A short physical performance battery assessing lower extremity function: Association with self-reported disability and prediction of mortality and nursing home admission. *Journal of Gerontology*, *49*, M85–M94. doi:10.1093/geronj/49.2.M85
- Hayes, R., & Moulton, L. (2009). *Cluster randomized trials*. Boca Raton, FL: CRC Press.
- Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, *32*, 341–370. doi:10.3102/1076998606298043
- Herrera, A. P., Meeks, T. W., Dawes, S. E., Hernandez, D. M., Thompson, W. K., Sommerfeld, D. H., ... Jeste, D. V. (2011). Emotional and cognitive health correlates of leisure activities in older Latino and Caucasian women. *Psychology, Health & Medicine*, *16*, 661–674. doi:10.1080/13548506.2011.555773
- Hodes, R. J., Insel, T. R., & Landis, S. C.; NIH Blueprint for Neuroscience Research. (2013). The NIH toolbox: Setting a standard for biomedical research. *Neurology*, *80* (11 Suppl. 3), S1. doi:10.1212/WNL.0b013e3182872e90
- Johnson, J. K., Gregorich, S. E., Acree, M., Napoles, A. M., Flatt, J. D., Pounds, D., ... Stewart, A. (2017). Recruitment and baseline characteristics of the Community of Voices choir study to promote health and well-being of diverse older adults. *Contemporary Clinical Trials Communications*, *8*, 106–113. doi:10.1016/j.conctc.2017.09.006
- Johnson, J. K., Louhivuori, J., & Siljander, E. (2017). Comparison of well-being of older adult choir singers and the general population in Finland: A case-control study. *Musicae Scientiae*, *21*, 178–194. doi:10.1177/1029864916644486
- Johnson, J. K., Napoles, A. M., Stewart, A. L., Max, W. B., Santoyo-Olsson, J., Freyre, R., ... Gregorich, S. E. (2015). Study protocol for a cluster randomized trial of the Community of Voices choir intervention to promote the health and well-being of diverse older adults. *BMC Public Health*, *15*, 1049. doi:10.1186/s12889-015-2395-9
- Kane, R. L., Butler, M., Fink, H. A., Brasure, M., Davila, H., Desai, P., ... Barclay, T. (2017). *Interventions to prevent age-related cognitive decline, mild cognitive impairment, and clinical Alzheimer's-type dementia*. Rockville, MD: Agency for Healthcare Research and Quality (US).
- Karp, A., Paillard-Borg, S., Wang, H. X., Silverstein, M., Winblad, B., & Fratiglioni, L. (2006). Mental, physical and social components in leisure activities equally contribute to decrease dementia risk. *Dementia and Geriatric Cognitive Disorders*, *21*, 65–73. doi:10.1159/000089919
- Kawagoe, T., Onoda, K., & Yamaguchi, S. (2017). Apathy and executive function in healthy elderly-resting state fMRI study. *Frontiers in Aging Neuroscience*, *9*, 124. doi:10.3389/fnagi.2017.00124
- Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B. W., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, *114*, 163–173. doi:10.1016/j.jad.2008.06.026
- Lorig, K., Stewart, A. L., Ritter, P., Gonzalez, V., Laurent, D., & Lynch, J. (1996). *Outcome measures for health education and other health care interventions*. Thousand Oaks, CA: Sage Publications.
- Louie, G. H., & Ward, M. M. (2011). Socioeconomic and ethnic differences in disease burden and disparities in physical function in older adults. *American Journal of Public Health*, *101*, 1322–1329. doi:10.2105/AJPH.2010.199455
- Macdonald, R., Kreutz, G., & Mitchell, L. (2012). *Music, health, & wellbeing*. Oxford: Oxford University Press.
- Marquine, M. J., Segawa, E., Wilson, R. S., Bennett, D. A., & Barnes, L. L. (2012). Association between cognitive activity and cognitive function in older Hispanics. *Journal of the International Neuropsychological Society*, *18*, 1041–1051. doi:10.1017/S135561771200080X
- Menec, V. H. (2003). The relation between everyday activities and successful aging: A 6-year longitudinal study. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *58*, S74–S82. doi:10.1093/geronb/58.2
- Mezuk, B., Choi, M., DeSantis, A. S., Rapp, S. R., Diez Roux, A. V., & Seeman, T. (2016). Loneliness, depression, and inflammation: Evidence from the multi-ethnic study of atherosclerosis. *PLoS One*, *11*, e0158056. doi:10.1371/journal.pone.0158056
- Napoles, A. M., & Chadiha, L. A. (2011). Advancing the science of recruitment and retention of ethnically diverse populations. *Gerontologist*, *51* (Suppl. 1), S142–S146. doi:10.1093/geront/gnr019
- Napoles, A. M., Santoyo-Olsson, J., & Stewart, A. L. (2013). Methods for translating evidence-based behavioral interventions for health-disparity communities. *Preventing Chronic Disease*, *10*, 1–12. doi:10.5888/Pcd10.130133
- Ngandu, T., Lehtisalo, J., Solomon, A., Levalahti, E., Ahtiluoto, S., Antikainen, R., ... Kivipelto, M. (2015). A 2 year multi-domain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): A randomised controlled trial. *Lancet*, *385*, 2255–2263. doi:10.1016/S0140-6736(15)60461-5
- Noice, T., Noice, H., & Kramer, A. F. (2014). Participatory arts for older adults: A review of benefits and challenges. *Gerontologist*, *54*, 741–753. doi:10.1093/geront/gnt138
- Pearce, E., Launay, J., Machin, A., & Dunbar, R. I. (2016). Is group singing special? Health, well-being and social bonds in community-based adult education classes. *Journal of Community & Applied Social Psychology*, *26*, 518–533. doi:10.1002/casp.2278
- Peplau, L. A. (1985). Loneliness research: Basic concepts and findings. In I. G. Sarason & B. R. Sarason (Eds), *Social support:*

- Theory, research and applications* (pp. 269–286). Boston, MA: Martinus Nijhoff.
- Perissinotto, C. M., Cenzer, I. S., & Covinsky, K. E. (2012). Loneliness in older persons : A predictor of functional decline and death. *Archives of Internal Medicine*, *172*, 1078–1083. doi:10.1001/archinternmed.2012.1993
- Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal*, *316*, 1236–1238. doi:10.1136/bmj.316.7139.1236
- Reitan, R. M. (1958). Validity of the trailmaking test as an indication of organic brain damage. *Perceptual and Motor Skills*, *8*, 271–276. doi:10.2466/PMS.8.7.271-276
- Reuben, D. B., Magasi, S., McCreath, H. E., Bohannon, R. W., Wang, Y. C., Bubela, D. J., ... Gershon, R. C. (2013). Motor assessment using the NIH toolbox. *Neurology*, *80* (11 Suppl. 3), S65–S75. doi:10.1212/WNL.0b013e3182872e01
- Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology*, *1*, 43–46. doi:10.1016/j.athoracsur.2015.11.024
- Sanchez-Cubillo, I., Perianez, J. A., Adrover-Roig, D., Rodriguez-Sanchez, J. M., Rios-Lago, M., Tirapu, J., & Barcelo, F. (2009). Construct validity of the trail making test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society*, *15*, 438–450. doi:10.1017/S1355617709090626
- Santoyo-Olsson, J., Cabrera, J., Freyre, R., Grossman, M., Alvarez, N., Mathur, D., ... Stewart, A. L. (2011). An innovative multiphased strategy to recruit underserved adults into a randomized trial of a community-based diabetes risk reduction program. *Gerontologist*, *51* (Suppl. 1), S82–S93. doi:10.1093/geront/gnr026
- Schneider, C. E., Hunter, E. G., & Bardach, S. H. (2018). Potential cognitive benefits from playing music among cognitively intact older adults: A scoping review. *Journal of Applied Gerontology*, 733464817751198. (Epub ahead of print). doi:10.1177/0733464817751198
- Smyth, N., Siriwardhana, C., Hotopf, M., & Hatch, S. L. (2015). Social networks, social support and psychiatric symptoms: Social determinants and associations within a multicultural community population. *Social Psychiatry and Psychiatric Epidemiology*, *50*, 1111–1120. doi:10.1007/s00127-014-0943-8
- Stern, C., & Munn, Z. (2010). Cognitive leisure activities and their role in preventing dementia: A systematic review. *International Journal of Evidence-Based Healthcare*, *8*, 2–17. doi:10.1111/j.1744-1609.2010.00150.x
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York: Oxford University Press.
- Tombaugh, T. N. (2004). Trail making test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology*, *19*, 203–214. doi:10.1016/S0887-6177(03)00039-8
- Trainor, L. J., & Hannon, E. E. (2013). Musical development. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 423–498). London: Elsevier.
- U.S. Census Bureau. (2017). *Profile America facts for features*. Washington, DC: U.S. Census Bureau.
- Varma, V. R., Tan, E. J., Gross, A. L., Harris, G., Romani, W., Fried, L. P., ... Carlson, M. C. (2016). Effect of community volunteering on physical activity: A randomized controlled trial. *American Journal of Preventive Medicine*, *50*, 106–110. doi:10.1016/j.amepre.2015.06.015
- Verghese, J., Lipton, R. B., Katz, M. J., Hall, C. B., Derby, C. A., Kuslansky, G., ... Buschke, H. (2003). Leisure activities and the risk of dementia in the elderly. *New England Journal of Medicine*, *348*, 2508–2516. doi:10.1056/NEJMoa022252
- Vincent, G. K., & Velkoff, V. A. (2010). *The next four decades. The older population in the United States: 2010 to 2050*. Washington, DC: US Census Bureau.
- Wilson, R. S., Bennett, D. A., Bienias, J. L., Mendes de Leon, C. F., Morris, M. C., & Evans, D. A. (2003). Cognitive activity and cognitive decline in a biracial community population. *Neurology*, *61*, 812–816. doi:10.1212/01.WNL.0000083989.44027.05