

A Comparison of the Effects of Short-term Singing, Exercise, and Discussion Group Activities on the Emotional State and Social Connectedness of Older Australians

Music & Science
Volume 1: 1–18

© The Author(s) 2018

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/2059204318800607

journals.sagepub.com/home/mns

Susan Maury and Nikki Rickard

Abstract

Choir membership has been shown to improve emotional states and facilitate social connectedness. It is, however, less clear whether these benefits are unique to group singing or are shared by other social group activities that include some of the characteristics of choirs other than singing, such as music listening and social interaction. This research compares older Australians who are members of either a choir that both produces and listens to music in a social context, an exercise group that incorporates music listening and movement with social interaction, or a current events discussion group with social interaction but no music content. Participants were administered emotional state and cohesion questionnaires at two test times, just prior to and immediately after the session, to determine the short-term (60–90 minutes) effects on emotional state and social cohesion as result of different social activities containing varying levels of music engagement. A two-way Analysis of Variance (ANOVA) revealed significant improvements in positive affect and cohesion scores, and a decrease in negative affect and tiredness scores, over time for all groups. The choir and exercise groups were also observed by two raters who recorded observable behaviors categorized using the circumplex model of emotion. Findings revealed that both groups demonstrated significant increases in Activated Pleasant (high positive affect, high arousal) behaviors over time, but with no differences between the two groups. Taken together, these studies suggest that well-being benefits are shared by self-selected leisure social group activities, and that the effects can be observed within a very short time frame using both self-report and behavioral measures. The authors suggest that future research incorporates suitable control groups into research designs to better articulate any unique benefits that group singing may confer.

Keywords

Music, choir, exercise, emotion, social cohesion, aging

Submission date: 24 October 2017; Acceptance date: 23 August 2018

Introduction

Social connectedness is a primary protective factor for healthy physical and mental aging. Loneliness in older years is associated with increased mortality, decreased physical capability and poorer mental health (Holt-Lunstad, Smith, Baker, Harris, & Stephenson, 2015; Luanaigh & Lawlor, 2008; Luo, Hawkey, Waite, & Cacioppo, 2012). While older populations have fewer social connections, they also have more time to invest in relationships and leisure activities (Cornwell, Laumann, &

Schumm, 2008). Previous research has identified a range of positive outcomes for older populations who engage in social activities, ranging from improved emotional

Monash University, Australia

Corresponding author:

Susan Maury, Monash University, 18 Innovation Walk, Clayton, Victoria 3800, Australia.

Email: Susan.Maury@Monash.edu



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons

Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial

use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on

the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

wellbeing and mental health (Horowitz & Vanner, 2010; Menec, 2003; Potočnik & Sonnentag, 2013), physical capability (Menec, 2003; Unger, McAvay, Bruce, Berkman, & Seeman, 1999) and cognitive functioning (Bennett, Schneider, Tang, Arnold, & Wilson, 2006; Seeman, Lusignolo, Albert, & Berkman, 2001). Importantly, supportive social connections with non-familial contacts have also yielded benefits (Fiori, Antonucci, & Cortina, 2006; Seeman, 2000), which highlights the role of broader community social opportunities for aging well. There has therefore been an interest in leisure activities for older populations which promote social connections as a means of maintaining mental health and wellbeing.

Group singing is a popular leisure activity which promotes social connectedness in several ways. It has been proposed that, from an evolutionary perspective, it creates a shared positive emotional state. Because emotions serve to focus attention on priorities for action (Cosmides & Tooby, 2000; Lang & Bradley, 2010; Lang & Davis, 2006), group-held emotions support the long-term viability of the group itself as well as commitment to jointly-held long-term plans (Niedenthal & Brauer, 2012). Thus, group singing may create a shared emotional state that facilitates both cohesion and cooperation (Cross, 2009; Maury & Rickard, 2016). Additionally, there is strong evidence that emotional musical experiences incorporate an empathic response (Cross, Laurence, & Rabinowitch, 2010; Egermann & McAdams, 2013; Miu & Balteş, 2012). A shared emotional state resulting from joint music-making may therefore encourage empathic responses to others. Increased empathy has strong links to increased prosocial behaviors (Eisenberg, Eggum, & Di Giunta, 2010; Eisenberg & Miller, 1987; Telle & Pfister, 2015), which reinforce group bonds and cohesion.

Groups which engage in music co-production have been shown to display both short-term changes in emotion and social connection (Clements-Cortés, 2015; Sanal & Gorsev, 2014) and more persistent changes (Bailey & Davidson, 2005, 2013; Dingle, Brander, Ballantyne, & Baker, 2013). While other types of social groups may also experience these benefits, there are indications that the pathway to these changes may be unique to music-producing groups (Pearce, Launay, & Dunbar, 2015; Pearce, Launay, MacCarron, & Dunbar, 2017). Studies which explore the short-term effects of music co-production may be able to perceive these differences, leading to a greater understanding of the mechanisms behind more persistent effects.

Group singing, affect and social bonds

Emotion regulation and social connection are often listed as the primary contributors to wellbeing for choir members. For example, increases in positive affect (PA) and social support were two mechanisms identified by choir members surveyed by Clift and Hancox (2010) (the others were

focused attention, deep breathing, cognitive stimulation, and regular commitment). von Lob, Camic, and Clift (2010) interviewed people who asserted choir membership has assisted them in coping with adverse life events; they named the collective experience of the choir and building relationships with other members to be important interpersonal mechanisms. When comparing a choir of homeless men to middle-class trained choristers, Bailey and Davidson (2005) found that emotional benefits were similar across the choirs, but that camaraderie was a particularly important aspect for the homeless choir members. Further, the same authors report that members of the homeless choir identified adaptive, persistent changes in emotion state and social skills resulting from choir membership (Bailey & Davidson, 2013). A choir formed for people with a disability provided members with important social benefits, including connecting with other people within the choir, connecting with the audience at performances, and improving overall social functioning (Dingle et al., 2013). People with clinical mental health issues in the UK experienced improved mental health over 8 months of choir membership, identifying emotional and social benefits as the mechanisms behind the improvement (Clift & Morrison, 2011). Welch, Himonides, Saunders, Papageorgi, and Sarazin (2014) reported that schoolchildren in a school-based singing program ($n = 6,087$) had a higher sense of self-concept and social inclusion, although this may be attributable to children developing a sense of mastery since self-concept and social inclusion were also positively correlated with children's level of singing ability.

Studies that have examined the short-term effects of group singing on mood have all reported an increase in PA, with mixed results for decreases in negative affect (NA). For example, when measuring experiences of flow amongst 44 students at a music academy, Fritz and Avsec (2007) found that those who performed in a choir or orchestra were the most likely to experience flow, and that flow states were positively associated with high subjective wellbeing, higher levels of PA and lower levels of NA. Small increases in PA and small decreases in NA were reported following a 1-hour choir practice with university students (Sanal & Gorsev, 2014) and with cancer patients or carers (Fancourt et al., 2016). Others report significant increases in PA but negligible changes in NA (Sandgren, 2009), including for people with Parkinson's Disease (Abell, Baird, & Chalmers, 2017). In a study with carers and older adults experiencing dementia or cognitive impairment, choir participation increased both PA and energy levels, while experiences of pain fell from pre- to post-session, and also across the 16 weeks of the experiment (Clements-Cortés, 2015).

The research reviewed above, therefore, supports that group singing is a leisure activity that is likely to have a positive influence on emotion state and social bonding. However, it is less clear whether choir membership is *more* likely to provide these benefits than other leisure activities.

Few studies provide a comparison group, and fewer still use a comparison group that provides a strong control in terms of similarity to the group singing experience. The following section reviews comparisons between group singing and other, non-music activities, followed by comparisons of differing types of music interventions.

Comparison studies

There are some studies which compare group singing with other, non-musical activities. Johnson, Louhivuori, and Siljander (2017) compared older (60–93 years) Finnish choir members ($n = 109$) with a demographically-matched sample ($n = 1,296$), some of whom were actively engaged in unspecified hobbies and others not. A comparison of quality of life (QoL) scores indicated that choir members reported significantly higher-rated overall QoL as well as health satisfaction compared with both control groups. Kreutz (2014) reports significant rises in PA and drops in NA as well as increases in oxytocin, a hormone implicated in the human bonding process, when comparing a singing to a chatting condition. Allpress, Clift, and Legg (2012) conducted a small ($n = 18$) study in which two groups participated in a 1-hour joint singing session or a 1-hour team exercise using Lego building blocks (R. Allpress, personal communication, 15 January 2014); the following day, the groups were switched. Differences reported on the two activities were not significant, although on both days the choir participants recorded slightly higher levels of PA and social cohesion and slightly lower levels of NA. Finally, an innovative study with young children found that playing a game resulted in increased cooperation and pro-social behaviors when the game included joint music-making (Kirschner & Tomasello, 2010).

At least two comparison studies report no differences between groups. For example, Dingle, Williams, Jetten, and Welch (2017) found no difference in affective changes between members of a choir and those enrolled in a creative writing class; this study also compared a normal older population with socially marginalized individuals, with no differences in mood changes between the groups. Similarly, a choir intervention with school children aged 7–11 years found no significant differences between choir participants and non-participants, although it is worth noting that the control group was small ($N = 10$ compared with 50 in the choir intervention) (Hinshaw, Clift, Hulbert, & Camic, 2015).

Recent research points to possible differences between how musical and non-musical groups achieve these benefits. Pearce et al. (2015) found that members of singing groups reported a stronger sense of social bonding after meeting regularly for 3 months compared with crafts and creative writing groups, although self-reports were identical between all groups after 7 months. Exploring these differences further, the authors found members of both the choir and creative writing conditions demonstrated higher

levels of person-to-person bonding than the crafts class, but the singers were the only group to develop a bond to the group as an entity (Pearce et al., 2017). These findings reinforce the need for examining wellbeing changes as they occur over a short period of time for music groups, in order to better understand how longer-term changes are achieved. Incorporating temporally sensitive methodologies into research – such as observations of behaviors over time (Bartel & Saavedra, 2000) or experience sampling methodology (Csikszentmihalyi & Larson, 2014) – may therefore illuminate subtle differences experienced in shorter time frames.

Few researchers have used a strong comparison group, such as those involving physical activity, exercise or sports. This is partly because “high exertive activity has been shown to increase social bonding” (Pearce et al., 2017, p. 498), which may be deemed to too closely replicate one of the key means by which group singing is likely to benefit its members. Alternatively, group exercise classes often include background music, which again dilutes the differentiation from the group singing intervention. A review conducted by Karageorghis and Priest (2012) reports that, when used with exercise, pre-task music can elevate arousal, and that self-selected music that is considered both motivating and stimulating can improve mood, reduce a sense of exertion, improve energy levels, and increase length of workout. It is to be expected, then, that the presence of music, provided it is pleasant to group members, would improve mood in both choir and exercise settings.

Stewart and Lonsdale (2016) provide an insightful study, comparing choral members to both solo singers and members of sports teams. Both choir members and sports team members reported significantly higher psychological well-being than solo singers; additionally, both choir and team members reported high levels of social bonding, although choir members were more likely to describe these social bonds as more meaningful. The authors suggest these findings may point to the importance of the group experience for socio-emotional wellbeing rather than the act of singing.

In summary, while group singing conditions appear to have a positive impact on affective state and social connection, comparison studies provide mixed results. Results from the children’s game designed by Kirschner and Tomasello (2010) indicate that group singing may promote sociability in groups in unique ways. However, in other findings changes in affect or social connection between music and non-music activities, when reported, are not always significantly different. It may be that the experiences between the groups are not identical, particularly concerning social processes. In particular, reports from Pearce and colleagues (Pearce et al., 2015, 2017) indicate that the singing groups in their study both bonded more rapidly and created bonds to the group as an entity, distinct from the strength of individual relationships that were formed across the length of the study. Therefore, while all the groups that

participated in their study reported increases in social cohesion, both the rapidity and the type of bonds felt by the singing group were different from the others. This is supported by Stewart and Lonsdale (2016), who report that choir members described their group as more coherent and meaningful than those described by members of sport teams. It also appears that comparison groups are generally selected in order to highlight the unique experience of music co-production, but this has left a gap in understanding how more similar social groups may either converge with or diverge from these experiences.

Music engagement. Studies that compare different methods of musical engagement are helpful to determine whether there are differences in the effects on wellbeing, for example between reception (that is, listening to music) and production (that is, singing or playing). Background music can have a positive impact on social interactions, including increasing a sense of “liking” in initial meetings (Stratton & Zalanowski, 1984b), increasing verbal exchange in social settings (Dubé, Chebat, & Morin, 1995; Stratton & Zalanowski, 1984a), and increasing positive assessment of an individual during an initial meeting (Ortiz, 1997). While there is ample evidence that music listening impacts on wellbeing (Croom, 2015; Juslin & Sloboda, 2010; MacDonald, 2013), including for older adults (Groarke & Hogan, 2015; Laukka, 2007), it is also becoming clear that musical preference is a key component; the music must be liked. Salimpoor, Benovoy, Longo, Cooperstock, and Zatorre (2009) examined this specifically with people who listened to both self-selected pleasurable and neutral musical pieces. They found that mood was elevated, with accompanying physiological responses, only when *liked* music was played. Thompson, Schellenberg, and Husain (2001) also found that listening to Mozart improved performance on a spatial task, but only when the listener found the music pleasurable, thus increasing both PA and arousal.

Therefore, the benefits of music listening are linked to how positively engaged listeners are with the music, which incorporates heightened arousal. While there are numerous music listening studies that measure changes in arousal (Grewe, Kopiez, & Altenmüller, 2009; Grewe, Nagel, Kopiez, & Altenmüller, 2007; Guhn, Hamm, & Zentner, 2007; Hirokawa, 2004; Rickard, 2004; Salimpoor, Benovoy, Larcher, Dagher, & Zatorre, 2011) it has seldom been measured in settings where music is being actively produced. One study reports that arousal increases were greater for individuals singing or tapping along to music, while these conditions as well as playing on a keyboard also decreased tiredness compared with a listening condition (Lim, 2008), indicating that active music production increases arousal. Contrary to these results, however, Grewe, Kopiez, and Altenmüller (2009) report no differences in physiological responses in a passive vs. active (singing along) condition to familiar music. These studies were conducted with individuals; the authors are unaware

of studies which measure arousal in a group setting. This is another gap in the research, since elevations in arousal may signal engagement with the music, and appears to also facilitate changes in mood (Chanda & Levitin, 2013; Salimpoor, Benovoy, Longo, Cooperstock, & Zatorre, 2009; van der Zwaag, Westerink, & van den Broek, 2011).

While it is known that music listening assists with affect regulation and influences social bonding (Dubé et al., 1995; Ortiz, 1997; Ziv, Granot, Hai, Dassa, & Haimov, 2007), there are few studies directly addressing the question of whether music production may provide greater benefits than a listening condition alone. However, there are indications that this may be the case. For example, when Kreutz, Bongard, Rohrmann, Hodapp, and Grebe (2004) compared the same choir at different times on both a singing and a listening condition, members reported significant increases in PA for both conditions, while NA dropped significantly for the singing condition, but rose for listening. A comparison of listening and singing effects on 5-year-old Japanese children found that they drew for longer and the drawings were judged as higher quality and more creative after singing familiar songs compared with a listening condition. Differences were also reported in the listening condition, with more proficient drawings produced after listening to children’s songs than after listening to classical selections (Schellenberg, Nakata, Hunter, & Tamoto, 2007). Baird et al. (2015) reported that singing both familiar and unfamiliar songs improved PA significantly compared with a non-musical condition for people with Parkinson’s disease.

In one of the few studies designed to explore differences between active music creation and listening, Dunbar, Kaskatis, MacDonald, and Barra (2012) found that a drumming group exhibited greater pain tolerance than a listening condition, additionally suggesting that the drumming condition facilitated group bonding more so than the listening condition through endorphin release. It is worth noting, however, that the listening condition was passive listening (in an office environment) rather than active, preferred listening. This was followed up with an active listening condition (fast vs. slow tempo music), which registered no increase in pain tolerance. The researchers conducted the same experiment comparing a choral or instrumental group with a dancing condition and found that the music groups outperformed the dancing condition in pain tolerance. These findings point to active music production, rather than either listening or coordinated movement, as the pathway for increased social bonding as measured through changes in pain tolerance.

In summary, it may be that wellbeing effects for music differ between listening and production; this topic could be explored through comparing music production with listening conditions in order to control for the effects of production. It has been established that, for listening conditions, benefits are greater for preferred music. Therefore, engagement with the music may be a critical component of wellbeing effects. Music engagement is distinct from training or

Table 1. Putative mechanisms present in each group type.

	Choir	Exercise	Discussion
Social connection	✓	✓	✓
Emotion regulation	✓	✓	×
Music listening	✓	✓	×
Music production	✓	×	×

competence, referring rather to an individual's emotional or intellectual commitment to a task (Chin & Rickard, 2012). Musically speaking, engagement can be present in a listening condition or absent in a production condition.

Aims of the current study

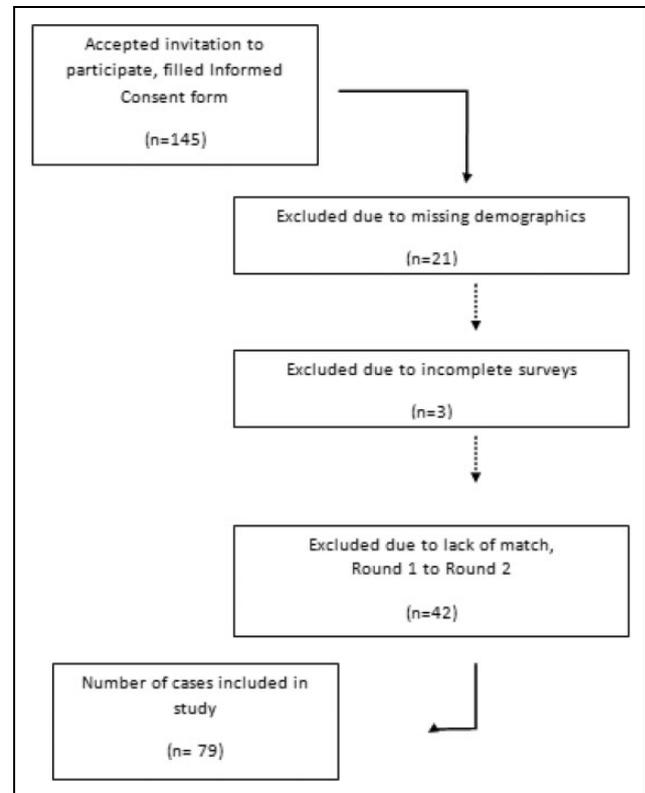
This study explores the effect of group-based leisure activities on social connectedness and emotion state measures of wellbeing for an older population in a natural setting. It seeks to differentiate whether effects are shared by each of the tested leisure groups, or whether there is a hierarchy of wellbeing effects depending on the level of music engagement: production, passive listening, or no music. To this end, non-auditioned choirs were compared with exercise groups which include music listening, and discussion groups that do not include music. This selection of groups was designed to provide insight into the potential mechanisms contributing to any benefits observed in the choir group, as several features of the group singing condition were shared by the control conditions (see Table 1).

In addition to changes in mood and social connection, this study also includes a measure of energy levels, since engagement with music has been shown to facilitate changes in arousal and may be linked to the emotional changes that individuals report. This current study is focused on short-term effects, pre- to post-session, as there are indications that the pathway to improved wellbeing may differ between musical and non-musical social groups. The short-term duration of the current study is an attempt to identify differences in experiences between the comparison groups that occur in the moment, over the course of a session.

Specifically, it was hypothesized that Choir members would report significantly greater increases in positive affect and greater reductions in negative affect pre- to post-session when compared with Exercise and Discussion groups. Choir members were also hypothesized to report similar increases in energy levels pre- to post-session when compared with Exercise groups, with both groups outperforming discussion groups. Finally, it was hypothesized that Choir members would indicate an increased sense of group bonding from pre- to post-session compared with both Exercise and Discussion Groups.

Method

This research included two components. Phase 1 involved completing a survey immediately before and immediately

**Figure 1.** Flowchart of participant attrition.

after a session, reporting on subjective mood, energy, and sense of social connection. Participants also provided basic demographic data. Phase 2 employed an observational matrix which was developed specifically for this research; more details on this are provided under *Phase 2 Materials and Procedure*. Both components of this research were approved by the Monash University Human Research Ethics Committee.

Phase 1

Participants. Participants from a sub-set of organized social groups from a larger study were approached to participate; see Figure 1 for flowchart of participant attrition. A power analysis was conducted by reviewing studies with a similar design, timeframe, and measures (Abell et al., 2017; Bartholomew & Miller, 2002; Ekkekakis, Backhouse, Gray, & Lind, 2008; Hirokawa, 2004; Kreutz, Bongard, Rohrmann, Hodapp, & Grebe, 2004; Lim, 2008). Based on the small to medium effects sizes reported in previous research ($f = .25$, $\alpha = .05$, and $\beta = .80$) and using a power analysis for a mixed measures Analysis of Variance (ANOVA) test, an estimated total sample size of 66, or 22 per group, was required for the short-term study design incorporating 3 groups across 2 timepoints.

Participants were sourced from several groups for each condition; three community non-auditioned choirs (Choir), four exercise (Exercise) groups, and two discussion (Discussion) groups. The exercise groups included medium- to

Table 2. Demographic frequencies, choir, exercise, and control groups.

	Choir		Exercise		Discussion	
	Frequency	%	Frequency	%	Frequency	%
Educational attainment						
No higher than year 10 of high school	2	7.7	9	33.3	3	11.5
Completed high school/VCE	2	7.7	1	3.7	2	7.7
Completed apprenticeship	0	0	0	0	1	3.8
TAFE/College diploma	8	30.8	5	18.5	4	15.4
Undergraduate university degree	6	23.1	5	18.5	5	19.2
Graduate diploma	2	7.7	4	14.8	1	3.8
Post graduate university degree	6	23.1	3	11.1	10	38.5
Employment status						
Unemployed (not studying)	3	11.5	2	7.4	0	0
Studying full time	0	0	1	3.7	0	0
Working part time	4	15.4	1	3.7	0	0
Working full time	4	15.4	0	0	0	0
Retired	15	57.7	21	77.8	26	100
Other			2	7.4		
Socio-Economic Standing (SES)						
Low SES	1	3.8	0	0	0	0
Average SES	18	69.2	24	92.3	19	73.1
Advantaged SES	7	26.9	2	7.4	4	15.4
Music Training (instrument or singing)						
Yes	18	69.2	9	33.3	10	38.5
No	8	30.8	18	66.7	16	61.5
Purposeful music listening						
Several hours/day	5	19.2	7	25.9	4	15.4
About one hour/day	8	30.8	5	18.5	6	23.1
Several times/week	9	34.6	7	25.9	11	42.3
Several times/month	4	15.4	4	14.8	3	11.5
Several times/year	0	0	4	14.8	2	7.7

VCE: Victorian Certificate of Education; TAFE: Technical and Further Education.

low-impact aerobics (listening predominantly to pop and “oldies” songs) and Tai Chi (listening to Chinese relaxation music). The discussion groups were both focused on current events. The majority of the groups were associated with the University of the Third Age (U3A) located in the outskirts of Melbourne, Australia. U3A is a social organization for people aged 55+ years which runs a wide range of groups that are organized and run by the members themselves, on a voluntary basis. All of the exercise and discussion groups in this study were attached to U3A, as well as one of the choirs. To balance numbers, other community, non-auditioned choirs were recruited to the study; this resulted in some demographic differences between the groups. All participants were informed that the study was exploring the possible social and emotional benefits of belonging to social groups, were provided with a printed information sheet, and were told that participation was optional. All participants signed and returned an informed consent form.

The mean age for the Choir group ($N = 26$) was 65.73 (SD: 8.24) years (age range 42–77 years), and the male/

female ratio was 3:23. The mean age for the Exercise group ($N = 27$) was slightly older, at 74.08 (SD: 6.97) years (age range 59–90), and the male/female ratio was similar at 2:25. The mean age for the Discussion group ($N = 26$) was 74.39 (SD: 5.04) years (age range 66–84), with a male/female ratio of 6:7. A one-way ANOVA confirmed significant age differences between the groups, $F(2, 75) = 13.25, p < .001$, and post-hoc tests revealed that the Choir was significantly younger than both the Exercise and the Discussion groups. A chi-squared test confirmed the differences in the sex ratios between the Discussion and other groups were also significant, $\chi^2(df = 2) = 14.06, p = .001$. Attempts to eliminate the age differential by containing cases to ages 65–80 years were unsuccessful and resulted in reducing the number of cases in the Choir condition to an unacceptable level. Similarly, it was not possible to reduce the gender ratio difference via any method of matching. Therefore, results should be interpreted with these demographic differences in mind. Additional demographic frequencies are displayed in Table 2.

The Discussion group appeared to be slightly more educated than the Choir or Exercise groups, but a chi-square

analysis revealed no significant difference in cell frequencies across groups, $\chi^2(df=12) = 15.99, p = .192$. The Choir group appeared to be more engaged in work than the other two groups, and a chi-squared confirmed this was significant: $\chi^2(df=10) = 25.02, p = .005$. Socio-economic standing (SES) differences were not significant across the groups, $\chi^2(df=4) = 5.80, p = .215$. Of those in the Choir group who had music training ($N = 18$), mean years of training were 4.72 (SD = 6.00); those in the Exercise group with music training ($N = 11$) had a higher mean of 6.18 years of music training (SD = 4.69), while those in the Discussion group ($N = 10$) had a mean of 5.11 years of training (SD = 3.44); no significant differences emerged, $F(2,35) = 0.28, p = .76$.

Table 2 shows the number of trained musicians in each group, and as expected, the Choir group had a significantly higher number of trained musicians, $\chi^2(df=2) = 6.41, p = .041$. Of those who indicated they had musical training across the three groups, the Choir group members practiced an instrument or sang between 1–5 hours/day (mean = 1.37, SD = 1.04) at the peak of their interest, while the Exercise group members with musical training practiced between 1–4 hours/day (mean = 1.35, SD = 1.00), and the Discussion group members practiced between 1–8 hours/day (mean = 2.33, SD = 2.43), with no significant differences emerging, $F(2,35) = 1.49, p = .239$. Frequencies of deliberate listening to music are displayed in Table 2; a chi-squared test revealed no significant differences across the groups, $\chi^2(df=8) = 6.62, p = .578$.

Materials

Demographic information collected included gender, age, postcode (to estimate SES), primary language spoken, handedness, education level, and employment status. Respondents were asked to list other organized social groups with which they were active. Three questions were included to measure music training: “Have you played/do you play a music instrument (includes singing, practice and performance)?”, “At the peak of your interest, how many estimated hours per day did you play/practice this primary music instrument (includes singing)?” and “How many years of musical training have you had?” There was also one question to serve as a proxy measure of music engagement; individuals were asked to estimate “On average, how often do you purposely listen to music a day (rather than to music in the environment that you have no control over, e.g., music in cafes, stores)?”

Self-report mood states were measured by the Positive and Negative Affect Schedule (PANAS) (Watson, Clark, & Tellegen, 1988). This widely-used measure of self-report positive and negative affect states was chosen for its reliability as well as its brevity, making it a good choice for the pre- and post-session design. It has been used in similar settings, including studies with choirs (Kreutz et al., 2004) and exercise groups (Bartholomew & Miller, 2002).

Participants were prompted to rate a list of 20 adjectives representing mood states on a scale of 1 (very slightly/not at all) to 5 (extremely) to the extent that they were feeling them this way right now, at the present moment. PA is measured by such adjectives as “Interested,” “Enthusiastic,” and “Attentive” whereas NA is measured by adjectives such as “Distressed,” “Upset,” and “Ashamed.” Cronbach’s alpha is reported at .89 for the PA scale, and .85 for the NA scale, with test–retest reliability reported as .79 (PA) and .81 (NA). The two scales are independently analyzed to provide a measure of both PA and NA. In the current study, Cronbach’s alpha indicated strong reliability at .94 for PA and .89 for NA (Time 1).

The Short Form Activation-Deactivation Checklist (AD ACL) (Thayer, 1978, 1986) was used to measure changes in energy levels. Similarly to the PANAS, the AD ACL is a widely-used measure that is very quick to complete and therefore was a good fit for the design of this study. The AD ACL has been used in music studies (Hirokawa, 2004; Lim, 2008) and in exercise settings (Ekkekakis et al., 2008). Participants are prompted to rate themselves on a list of 20 adjectives which describe how active and energetic they are feeling right now, on a scale of 1 (definitely do not feel) to 4 (definitely feel). The AD ACL includes four sub-scales: Descriptors for the Energy sub-scale include “Active,” “Energetic,” and “Full of pep”; examples for the Calmness sub-scale include “Still,” “Quiet,” and “Placid”; Tiredness adjectives include “Wide awake” (reverse scored), “Drowsy,” and “Sleepy”; the Tension sub-scale includes “Jittery,” “Fearful,” and “Clutched up.” Test–retest reliability ranges between .75–.92, while alphas for the subscales range between .89–.92 (Thayer, 1978). Sub-scales were shown to have acceptable reliability, with Time 1 Cronbach’s alphas of .92 (Energy), .85 (Calmness), .74 (Tension), and .85 (Tiredness).

A search failed to find a questionnaire designed to measure short-term changes in group cohesion. The Measures of Psychological Climate, Cohesion sub-scale (Koys & DeCotiis, 1991) was adapted for the current study. Published alphas range between .82–.95. Results for the current study demonstrated good internal reliability, with Cronbach’s alphas of .9 (Time 1). The sub-scale includes 5 statements with a 7-point scale ranging from 1 (Completely Disagree) to 7 (Strongly Agree); participants were asked to indicate how they feel the group interacts right now. As this scale was originally used to measure group cohesion within a work environment, the 5 statements were slightly modified to refer to a “group” rather than an “organization”; examples of statements include “In this group, people pitch in to help each other out,” and “there is a lot of ‘team spirit’ amongst this group.”

Procedure

Participants were tested in the natural environment of their regular session. Groups ranged in size between

Table 3. Results on Phase I pre- and post-tests measures, all groups, with time and time \times group interaction p value.

Measure/ Sub-scale	Choir M (SD)	Choir T2 M (SD)	Exercise M (SD)	Exercise		Discussion		Time p value	Effect size (Cohen's d)	Time \times Group Interaction p value	Effect size (Cohen's d)
				T2 M (SD)	M (SD)	T2 M (SD)	M (SD)				
PANAS											
PA	3.33 (0.92)	3.7 (0.85)	3.38 (0.86)	3.66 (0.84)	3.08 (0.95)	3.27 (0.85)	.000 ^a	1.04	.187		0.29
NA	1.22 (0.53)	1.07 (0.26)	1.20 (0.31)	1.09 (0.26)	1.27 (0.42)	1.27 (0.51)	.074	.57	.270		0.38
AD ACL											
Energy	3.08 (0.77)	3.22 (0.75)	3.25 (0.76)	3.37 (0.67)	2.86 (0.86)	2.89 (0.68)	.152	.33	.784		0.16
Calmness	2.65 (0.59)	2.38 (0.59)	2.35 (0.69)	2.39 (0.73)	2.17 (0.97)	2.19 (0.67)	.423	.18	.302		0.36
Tension	1.38 (0.53)	1.25 (0.42)	1.37 (0.54)	1.29 (0.45)	1.42 (0.49)	1.49 (0.70)	.268	.255	.220		0.40
Tiredness	2.10 (0.74)	2.03 (0.63)	2.06 (0.75)	1.76 (0.48)	1.93 (0.63)	1.76 (0.37)	.012 ^a	.059	.382		0.32
Group Cohesion	5.88 (0.78)	6.10 (0.72)	5.81 (0.80)	5.98 (0.73)	5.48 (1.01)	5.46 (1.05)	.047 ^a	.46	.236		0.39

^aThe difference in mean values is considered significant at the .05 level.

35–60 people, and for each group between 40–90% agreed to participate in the study. A survey was administered by the researcher just prior to the session, and again immediately following the session. All of the Exercise sessions and two of the Choir sessions ($n = 10$) ran for a duration of 60 minutes. The remaining choir ($n = 16$) ran for a duration of 90 minutes. A two-way mixed measures ANOVA was performed on the Choir groups to check whether the differing length of sessions had an effect on outcomes. Results showed a significant difference for changes in Energy levels, with the 60-minute Choirs reporting a small decrease in energy levels, and the 90-minute choir reporting an increase, $M_{diff} = 0.29$, $F(1,22) = 8.67$, $p = .008$. No differences arose with any of the other measures. Each of the eight individual groups (three Choir, three Exercise, two Discussion groups) was led by a different facilitator and held in different meeting spaces, at differing times of day (most were held in the morning, but some met in the afternoon and one met in the evening). These unique aspects of delivery were across all group types. All groups were tested at the same time of year to control for seasonal affect changes. Pre- and post-surveys were identical excepting the demographic data, which was included in the post-session questionnaire in order to reduce disruption to the session, since its inclusion extended the survey's length.

Data analysis

All measures were analyzed by a two-way mixed measures ANOVA using SPSS version 24. Twenty-six cells were missing data and were replaced with the sub-scale mean. Three cases were outliers on the Tension subscale of the AD ACL and one on the NA subscale of the PANAS. Since these cases were within range on other measures, the decision was made to Winsorize these cases rather than trim them; this process adjusts the outliers to within 3.29 SD of the mean so that they are less extreme and less likely to skew results. Preliminary checks were conducted to ensure

that there was no violation of the assumptions of normality of distribution and homogeneity of variance. An alpha level of .05 was used for all statistical tests.

Results

Mean values on all self-reported measures pre- and post-intervention for each of the three groups are shown in Table 3.

The 2×3 mixed ANOVAs yielded no significant interaction effects. Despite this lack of statistical significance, a consistent trend was that the Choir group reported slightly greater increases in PA and Cohesion over time than did either the Exercise or Discussion groups. Similarly, the Choir group reported slightly greater decreases in negative outcome measures (i.e., NA and tension) than did the other two groups. The Exercise and Discussion Groups both reported experiencing greater decreases in the Tiredness measure than the Choir Group.

Main effects for time were significant on the measures of PA, NA, Tiredness, and Cohesion. When averaged across groups, PA ratings rose significantly across the two time points, $F(2,76) = 20.48$, $p = .001$, 95% confidence interval (CI) $[-.40, -.16]$, as did Cohesion ratings: $F(2,76) = 4.06$, $p = .047$, 95% CI $[-.25, -.001]$. NA ratings decreased over time, $F(2,76) = 6.12$, $p = .016$, 95% CI $[-.02, .16]$, as did Tiredness: $F(2,76) = 6.68$, $p = .012$, 95% CI $[-.04, .32]$. There were no main effects of time for Energy, Calmness, or Tension.

Phase 2

The second component of this research utilized an observational methodology for the Choir and Exercise groups only. It was reasoned that if group members were experiencing changes in emotional state, energy levels, and a sense of group cohesion, changes in behavior should reflect this. This study was designed to complement Phase 1, and also

Table 4. Brief description of observed groups.

Group type	Group name	Description	Approx. # of people observed	Frequency and # of observations
Choir	University of the Third Age Choir	This no-audition choir is the only choir in the study associated with the University of the Third Age. It sings a range of mostly popular and musical songs from the 1950s – 1990s. Rehearsals are accompanied by a pianist.	T1*: 28 T2: N/A T3: 25	T1: 305 T2: 0 ^a T3: 282
Choir	Open Door Singers	This no-audition community choir meets in the evening for approximately one and a half hours. Song selection tends to be popular songs from a range of styles, and are sung off of an overhead projection, and accompanied by either the director on guitar or by a soundtrack.	T1: 28 T2: 40 T3: 15	T1: 378 T2: 424 T3: 153
Choir	Box Hill Choir	This community choir meets at a community arts center. Of the choirs involved in this study, this choir sang the most musically challenging pieces, primarily written for performance choirs and not in the popular canon.	T1: 10 T2: 15 T3: N/A	T1: 112 T2: 197 T3: 0 ^b
Exercise	Moderate Exercise	This group was reasonably fit and flexible. Workouts were done to mostly upbeat music, incorporating aerobics, country dancing, fast walking, and finishing with mat stretching to calming music and dimmed lights.	T1: 20 T2: N/A T3: 30	T1: 217 T2: 0 ^a T3: 71
Exercise	Tai Chi	This large class meets for 1 hour 30 minutes in the morning with a small break in the middle. The class follows a leader through gentle tai chi moves, accompanied by traditional Chinese music.	T1: 22 T2: 50 T3: 20	T1: 313 T2: 417 T3: 146

*T1 = Pre-session; T2 = mid-session; T3 = post-session.

^aNo break was taken.

^bMembers left immediately.

to test whether behavior changes could be reliably observed and recorded across a large group.

Participants

Phase 2 included the same Choir and Exercise groups that participated in Phase 1, although on a different day. The majority of observation sessions were conducted 2 weeks after the surveys, although for one group it was 2 months later due to scheduling difficulties. A researcher visited the group the week prior to explain the observation session, and then on the day individuals were invited to sign a consent form for being observed and were told if they wished not to be observed to identify themselves to the researchers. All participants provided consent for Study 2 observations. A total of three Exercise and three Choir groups were observed; however, one of the Exercise groups was eliminated from analysis as they started their session early and a baseline observation was therefore not captured. A brief description of each group is provided in Table 4. The approximate number of people observed is provided at each observation time point – that is, prior to the start of the session (Time 1), at the mid-point break if any (Time 2) and after the end of the session (Time 3). This methodology is unique in that it is capturing the number of observations of certain behaviors rather than the number of participants per se, so the number of participants is an estimate only. The number of people observed in a choir setting was approximately 95, while the number of people observed in an exercise class was approximately 80.

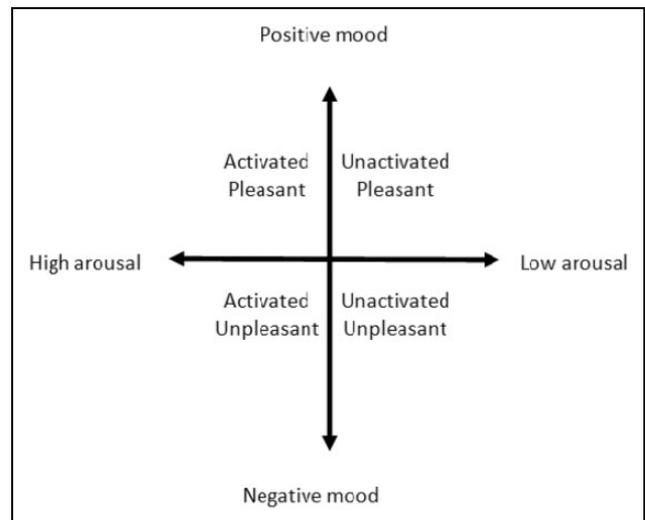


Figure 2. Circumplex model of emotion (Adapted from Russell, 1980).

Materials

This study used an observational checklist which was adapted from that created by Bartel and Saavedra (2000) for use within organizations. The checklist is based on the circumplex model of emotion (Russell, 1980), and identifies specific physical behaviors that can be categorized within one of four quadrants: Activated Pleasant (high arousal, high PA), Unactivated Pleasant (low arousal, high PA), Unactivated Unpleasant (low

Table 5. Observational checklist.

	Activated Pleasant	Unactivated Pleasant	Unactivated unpleasant	Activated unpleasant
Body	Leaning forward Orienting towards others	Relaxed but engaged orientation towards group	Orienting away from group Slouching	Body poised to exclude group members
Movement	Constant body movement	Little movement in torso or limbs	Motionless Resting head on hands	Nervous habits (rocking, biting fingernails)
Physical Contact	High physical contact	Moderate contact	No contact	Avoiding contact
Hands	Exaggerated hand gestures Hands active during speech	Minimal hand movement	Hands inactive during speech Rubbing eyes	Closed fists Hand tremors
Mouth	Smiling with teeth showing Grin (big closed lipped smile)	Mouth turned slightly upwards, open or closed	Yawning Mouth turned downwards	Sneering Clenched teeth
Eyes	High eye contact	Moderate eye contact	Little eye contact Blank stare	Avoiding eye contact

arousal, high NA), and Activated Unpleasant (high arousal, high NA); see Figure 2.

The full checklist with prompts is reproduced above (see Table 5). An eighth row for observing eyebrows was removed as this was too difficult to observe in a large group setting.

Procedure

A primary researcher and two research assistants trialed the tool ahead of time using both video and live observation; a one-page instruction sheet was also created to ensure as much as possible a uniform approach to coding. In each observation, two raters were used – the primary researcher and an assistant. In order to mitigate possible rater bias, the tool was designed to be as objective as possible by recording specific and visible changes in how the body or the face was positioned. The raters observed each group at up to three time points: prior to the start of the session, at the mid-point break (if there was one) and at the end of the session. Each rater recorded their name, the date and time, a short description of what the group was doing (e.g., waiting to start, having a snack), and an estimate of how many people were being observed. Observations were done simultaneously by both raters, scanning the room in the same direction (for example, from left to right) and from the same vantage point in order to provide congruent observations as much as was possible. Each rater observed each person in the room only once, recording all aspects of an individual's body language that could reliably be seen by recording a hash mark in the relevant box. Depending on the number and distribution of people in the room, this took between 2 and 5 minutes. Videotaping sessions was considered, which would have facilitated coding that was more uniform and also allowed for naïve raters to be used. However, there was concern that the taping would be disruptive to the groups, make individuals feel uncomfortable, and would have limited what was observable.

This methodology provided a tally of observations of individuals within the quadrants for Body – Movement – Physical Contact – Hands – Mouth – Eyes. However, not each quadrant was observed for each individual in the room. For example, people who had their back turned to the researchers could be observed for Body, Movement, Physical Contact, and possibly Hands, but not for Mouth or Eyes. In addition, each researcher made independent observations, so while efforts were made to keep observations as congruent as possible, it was not feasible to ensure both were observing the same person at the same time. Therefore, observations of the same person sometimes resulted in differences in the tallies. The same people were not necessarily viewed at the three timepoints, as some arrived late or left early. For example, Table 4 indicates that the fewest observations were taken at Time 3 (after the session ended). This was due to many leaving immediately at the conclusion of the session. Reasons for this were varied, but in some cases was because U3A members had to rush to attend their next class.

Due to the unique challenges of this data set, the group itself is treated as a unit, within which changes of behavior can be observed. It was therefore decided to aggregate the observation tallies from each rater and report the mean. In addition, because there was unequal opportunity for observations at each of the three timepoints, means were standardized across the timepoints as a percentage of the total number of observations for that session. Chi-squared analyses are reported on the changes in frequencies across time, but should be interpreted with caution for these reasons.

Results

Table 6 shows the percentage distribution of observed behaviors across the three timepoints for both the Choir and Exercise Groups, with the expected statistical distribution reported in parentheses. A chi-square goodness of fit analysis shows that both the Choir ($\chi^2 (df = 2) = 11.56$,

Table 6. Observed changes in mood and energy, Choir and Exercise groups observational data (expected values in brackets), with goodness of fit calculations across time for each emotion quadrant.

Emotion Quadrant	Time 1	Time 2	Time 3	Chi-square
Choir				
Activated Pleasant (high energy, positive mood)	17 ^a (27)	23 (27)	41 (27)	11.56, $p = .003^b$
Unactivated Pleasant (low energy, positive mood)	49 (46)	50 (46)	39 (46)	1.61, $p = .447$
Unactivated Unpleasant (low energy, negative mood)	33 (25)	24 (25)	18 (25)	4.56, $p = .102$
Activated Unpleasant (high energy, negative mood)	2 (2)	2 (2)	2 (2)	0, $p = 1.0$
Exercise				
Activated Pleasant (high energy, positive mood)	12 (24)	27 (24)	33 (24)	9.75, $p = .008^a$
Unactivated Pleasant (low energy, positive mood)	55 (46)	51 (46)	47 (46)	.627, $p = .731$
Unactivated Unpleasant (low energy, negative mood)	28 (21)	19 (21)	16 (21)	3.714, $p = .156$
Activated Unpleasant (high energy, negative mood)	4 (4)	3 (4)	5 (4)	.5, $p = .779$

^aThere were varying degrees of opportunity to observe across the three sessions, as detailed in Table 4. For example, some groups did not take a mid-session break, and many participants left quickly at the end of sessions to attend to other obligations. Therefore, the observations have been standardized as a proportion of total observations in any one sitting, to control for non-specific effects associated with times.

^bIndicates changes in this quadrant vary significantly from the expected distribution.

$p = .003$) and Exercise Groups ($\chi^2 (df = 2) = 9.75$, $p = .008$) experienced a significant increase of observed behaviors in the Activated Pleasant (positive mood, high energy) quadrant. Changes in the other three quadrants conformed to expected distribution patterns for both groups, with non-significant decreases in the Unactivated Pleasant and Unactivated Unpleasant quadrants and no appreciable change in the Unactivated Unpleasant quadrant.

Discussion

This study explored the short-term (pre- to post-session) effect of group-based leisure activities on social connectedness and emotional state measures of wellbeing, and whether these effects might relate to differing levels of music use. Based on previous research, it was anticipated that there may be a hierarchy of wellbeing effects depending on the level of music engagement: production, passive listening, or no music. To this end, non-auditioned choirs were compared with exercise groups which include music listening, and discussion groups that did not include music, with members drawn from an older population in a natural setting. The use of exercise groups provided a more rigorous comparison than is often used with choirs, since the combination of music listening, movement, and social interaction more closely aligns with non-specific features of choir groups, while the discussion group provided a comparison with a no-music, no-movement but still engaging group condition. By comparing with groups that share more of the non-

specific characteristics of a choir experience, it is possible to provide a more systematic test of the wellbeing benefits that are ascribed to group singing, to determine whether they are attributable to that activity per se.

An absence of evidence was obtained in support of the three hypotheses of this study. First, it was predicted that Choir members would report significantly higher increases in PA and reductions in NA pre- to post-session compared with the Exercise and Discussion Groups. This hypothesis was not supported by the self-report scores provided pre- and post-session by participants reported in phase 1, in which increases in PA and sense of group cohesion were experienced across all three groups. Further, while the Choir Group evinced a significant increase in positive mood and increased energy via observable behaviors in phase 2 of this study, these increases were similarly observed in the Exercise Group. The second hypothesis – that the Choir Group and Exercise Group would report similar increases in Energy pre- to post-session while the Discussion Group would not – was not supported. Although both the Choir and Exercise Groups reported small increases in Energy, these changes were not significantly different to those occurring in the Discussion Group. Finally, it was predicted that the Choir Group would report higher Cohesion ratings than both the Exercise and Discussion Groups from pre- to post-test, which was not supported. While the Choir Group rated higher on Cohesion than the comparison groups, the differences were not significant.

Previous research into the wellbeing effects of choir membership indicates that benefits ought to be greater when compared with other groups. More generally, the wellbeing benefits of music listening are well documented; including any kind of music component into a social group setting would generally be expected to increase wellbeing effects, provided music engagement is high (Chin & Rickard, 2012). It is therefore notable that no significant differences across the groups included in this study were observed on the measures of mood, arousal or social cohesion. The current findings raise the possibility that significant short-term benefits of choir participation observed in previous research may also be attributable to characteristics of the activity not specific to choirs, such as group engagement, self-selection, and perhaps music exposure, rather than music production per se. A review of previous research demonstrates this may indeed be the case. A number of studies which report significant benefits did not include a control group (e.g., Abell et al., 2017; Fancourt et al., 2016; Sandgren, 2009), so are unhelpful in exploring this possibility. Several studies which observed significant differences between a choir group and a comparison are open to alternative explanations. For example, it appears that the differences in PA ratings between a choir and control group observed by Sanal and Gorsev (2014) may be attributable to a *decrease* in PA for the control group (who had self-selected to be part of a choir but were prevented from singing on the test day). In contrast, PA remained stable in the choir group pre- to post-singing. Similarly, Pearce et al. (2015) reported increased PA for members of a choir intervention compared with non-singing group activities at a community center; however, these differences are the result of a lower self-report of PA by the choir members in the pre-session measures (Choir Time 1: $M = 2.93$, $SD = 0.97$; Choir Time 2: $M = 3.58$, $SD = 1.04$) than the control group (Control Time 1: $M = 3.5$, $SD = 0.99$; Control Time 2: $M = 3.50$, $SD = 0.99$), rather than a significantly higher report of PA at the end of class (3.58 for Singing compared with 3.50 for Control), so it is difficult to exclude external factors present in this group at baseline which may have impacted on PA. Pearce and colleagues also reported on sense of social connection through a self-report and a proxy measure of pain tolerance. While pain tolerance increased pre- to post-session, there were no differences between the groups, while self-reported closeness to the group pre- to post-session was higher for the choir group in two of the three testing sessions. Kreutz et al. (2004) and Kreutz (2014) report on a choir made up of members who had self-selected to join a choir. They were compared with themselves, one week in a singing condition, and one week in either a listening (Kreutz et al., 2004) or a chatting (Kreutz, 2014) condition. While both studies yielded significant differences between the two conditions, it may be that the differences are due to withholding a preferred activity rather than indicating that singing has increased benefits to either of the other activities per se.

In contrast, when appropriate controls have been used, the difference is not convincing. Allpress et al. (2012) recruited a naïve population and randomly allocated them to either a singing or Lego-building activity, with groups switching activities on day 2. No significant difference in measures of mood or cortisol were found between groups. Research conducted by Dingle et al. (2017) also found no differences in mood rating for members of a community choir compared with members of a range of arts-based groups specifically for people with compromised mental health on the day of their activity; importantly, however, this study was not comparing a singing with a non-singing condition, as the “arts-based groups” included both a choir and creative writing classes. Taken together, none of these studies demonstrates convincing evidence that a group singing session may confer more benefits to mood when compared with other kinds of self-selected group activities. The studies which examined short-term changes in group cohesion are fewer but more promising. As part of the study conducted by Kreutz (2014), discussed above, saliva samples were also analyzed for changes in cortisol, oxytocin, and DHEA at the start of each session and again 30 minutes later. Significant time \times condition changes were found in oxytocin levels – a biomarker of bonding with others – with the singing condition experiencing significantly higher levels from pre- to post-condition compared with the chatting condition. This is promising, since one may expect a chatting condition to increase a sense of bonding to others, particularly as this involved sharing personal stories that brought past happiness. However, this may again be explained by participating in a preferred activity compared with having the activity denied. Kirschner and Tomasello (2010) found that preschool children who played a game with a singing and movement component showed greater cooperative and helping behavior than those who played the same game without singing or movement. While this study presents the strongest evidence for a superior effect of a singing activity than no singing activity, it is notable that the control group was also denied movement, which may also have confounded interpretation.

The naturalistic setting maintained in the current study is a strength of this research. Assessing individuals in situ is likely to have achieved greater ecological validity of observed behaviors and may have increased accuracy of self-reports as individuals were not required to recall their emotional state in a different setting. Further, it was an opportunity to explore how wellbeing effects of various leisure activities are experienced in everyday life, where a combination of factors, including motivation and choice, interact to influence wellbeing in ways that are not possible to observe in a more highly controlled environment. However, the natural setting may also explain the absence of significant differences between the groups in this study. First, the groups that were included in this study had been established for some time, which means that relationships had already been formed. This likely reduced the impact of

the interventions on all measures, making changes in emotion and social connection more difficult to detect due to a “ceiling effect” of sorts. Second, the members of U3A are part of a larger network of social groups and classes. Most, if not all, of the U3A participants attend multiple sessions throughout the week, and often meet up for coffee or lunch in between sessions. Research indicates that strong social networks are highly protective for the wellbeing of their members (House, Landis, & Umberson, 1988; Steptoe, Shankar, Demakakos, & Wardle, 2013), and again may mean the effects of attending one session in the day may be blunted. Finally, there were differences between the eight individual leisure groups which participated in this research; these included the length of time for sessions (which ranged between 60–90 minutes), time of day, the size of the group, and the personality and leadership style of the leader. It was not possible to control for all of these differences but as these differences were evenly distributed across group types, sampling across several leisure groups may have helped to mitigate their influence on the outcomes.

It is also possible that the demographic differences between the groups influenced the outcomes, possibly masking what would otherwise have been measurable differences. The Discussion group had a more balanced ratio of males to females than the other two groups. It appears from previous research that the benefits of social groups are greater for women than for men (Agahi & Parker, 2008), as are the benefits received from choir participation (Sandgren, 2009). However, these differences would therefore be expected to increase rather than reduce differences between the Discussion group (closely balanced between men and women) and the Choir and Exercise groups (primarily women). The Choir group was also both younger and more engaged in work, while the Exercise and Discussion groups were made up almost exclusively of retired individuals. Amongst other differences, it is likely this resulted in inequalities in overall general health, mobility, and patterns of social engagement, which may have influenced findings. However, as with the differences in gender balance between the groups, this would be expected to magnify rather than diminish differences in wellbeing benefits between the groups. It could be too that, as participants were informed that the study was looking at possible socio-emotional benefits to social group membership, demand characteristics may have impacted more on the Choir groups, since the wellbeing benefits of choir membership are widely publicized.

It could be, however, that the subtle differences between the groups may be more pronounced for other cohorts or in other settings, as has been found in research with more marginalized populations. Although not reaching significance, the Choir group reported slightly higher changes in Cohesion ratings in the pre- to post-session self-report. Research conducted with vulnerable populations and over longer periods of time reports that social connection is a

primary benefit identified by choir participants, including choirs for people experiencing homelessness (Bailey & Davidson, 2005, 2013), disability (Bailey & Davidson, 2005, 2013; Dingle et al., 2013), compromised mental health (Clift & Morrison, 2011), and those experiencing adverse life events (Clift & Hancox, 2010; Fancourt et al., 2016; von Lob, Camic, & Clift, 2010). The sample of older adults in the study reported here are generally very well socially connected, as many were members of active social networks through their connection with U3A (members of the non-U3A-related community choirs also indicated in their demographic responses that they were active in a range of other organized social groups). Higher levels of social connectedness may mute any additional effects of choir participation on wellbeing. This aligns with findings reported by Hinshaw, Clift, Hulbert, and Camic (2015), in which young people participating in a choral intervention did not evidence an increase in psychological wellbeing compared with a control group, although the participants did describe a range of benefits from participating. The authors suggest that benefits may be difficult to measure for groups which are already experiencing high levels of wellbeing and social connection.

An alternative explanation is that any social group will confer similar wellbeing benefits to choir membership. It may be that the three groups conferred similar levels of wellbeing because members self-selected into group activities that were of particular interest to them. Self-selection may explain why even the Discussion group, which lacked both music and coordinated movement, still improved overall markers of wellbeing. As demonstrated in the review of short-term comparison studies, withholding of a preferred activity appears to reduce self-reported PA, while comparing across self-selected but differing activities does not yield appreciable differences between activities. This is consistent with Social Determination Theory (SDT), which identifies autonomy, competence, and relatedness (that is, a sense that individual actions contribute to the wellbeing of others and are appreciated) as basic psychological needs of all humans (Deci & Ryan, 2008; Ryan & Deci, 2000). Within this framework, the overriding principle of wellbeing would be an individual’s motivation and choice for selecting an activity, rather than the nature of the activity itself. This has also been proposed as a possible explanation for the benefits of singing by Stewart and Lonsdale (2016), and may need further exploration.

The research reported here makes an important contribution to our understanding of music’s effects on emotion regulation and social cohesion. The short timeframe of the pre- to post-session design and the observation sessions provide ways to track changes within the group as they happen. In addition, the observation methodology tests a way of recording observable changes in emotion and arousal across a large group. While no appreciable differences were found between the Choir and Exercise groups in this study, this trial of the methodology did confirm that

changes in group interactions are observable and can be tracked across time. It would be useful to trial this tool in other settings, perhaps using videotaped footage in order to mitigate possible rater bias. Finally, the use of exercise groups which incorporate music listening and movement provides a strong comparison group from which to gauge whether the positive impacts of singing groups on wellbeing are unique, or whether other groups which incorporate similar processes may deliver the same or greater benefits.

The findings from this study call into question the hypothesis that group singing provides greater support to improved mood and social connection compared with other kinds of group leisure activities within a short timeframe. A reappraisal of previous studies exploring short-term affective and social cohesion changes in a singing condition suggests that claims of any unique effect of singing may be overstated. Future research could seek to identify the key components of the choir experience that are primarily responsible for explaining how group singing improves wellbeing. Identifying the mechanisms that are responsible for wellbeing improvements could lead to a greater understanding of how other, non-musical interventions may provide the same benefits. For example, flow states, which reflect a balance of challenge and ability and is marked by rewarding and enjoyable absorption in an activity (Nakamura & Csikszentmihalyi, 2014) is often reported in music production (Chirico, Serino, Cipresso, Gaggioli, & Riva, 2015; Sinnamon, Moran, & O'Connell, 2012) and appears to be heightened in a group setting (Páez, Rimé, Basabe, Włodarczyk, & Zumeta, 2015; Walker, 2010). There is also a growing body of evidence that synchronous movement amongst groups both improves mood, facilitates group cohesion, and may support an increase in cooperation and pro-social behaviors (Valdesolo & DeSteno, 2011; Valdesolo, Ouyang, & DeSteno, 2010; Reddish, Bulbulia, & Fischer, 2014; Trainor & Cirelli, 2015). Further, it may be that, rather than either production or listening, musical engagement is the critical mechanism behind improved wellbeing. The concept of engagement is understudied (Chin & Rickard, 2012) and may explain the similarities between the Choir and Exercise Groups, since demographic responses indicate little difference between the two groups in musical engagement. Relatedly, a heightened empathic response is linked to music engagement and emotion contagion (Egermann & McAdams, 2013), and is another possible pathway for wellbeing benefits. Heightened empathy is correlated with an increase in pro-social behaviors, and there is concern that as adults age, they experience reductions in cognitive empathy (the theory-of-mind ability to infer the emotional state of another person), which may in turn erode their social interactions (Sze, Gyurak, Goodkind, & Levenson, 2012). Therefore, finding ways – employing music or other social activities – to retain empathic responses for older adults may assist them to maintain social networks. In the light of this study's findings,

however, more research into the role of preference, choice, and agency in selecting leisure group activities may be the most promising direction to pursue. SDT may be a helpful way to frame further research into the links between wellbeing and participation in leisure group activities.

Finally, this study focused on short-term changes across groups on socio-emotional wellbeing. A longer-term study may reveal wellbeing benefits that emerge only over an extended period of time. If music and non-musical groups do bond differently over time, there may also be differences in the way that individuals interact longer-term, both within the group itself as well as more generally. For example, if evolutionary theories are correct, we would predict that members of singing groups would display heightened empathy (Cross et al., 2010; Greenberg, Rentfrow, & Baron-Cohen, 2015) and pro-social behaviors (Koelsch, 2013; Schulkin & Raglan, 2014). There are indications that this may be so. For example, a systematic review indicates that singing programs can assist individuals with dementia to increase social behaviors, encourage participation, and reduce anxiety (Clift, Nicol, Raisbeck, Whitmore, & Morrison, 2010). A longitudinal study of older adults found that those who participated in a choir reported fewer doctor visits, reduced medication, fewer falls, and better overall health than those who participated in self-selected (unspecified) activities; choir members also trended towards increased social participation more generally while the non-choir group reduced their social participation over the same time period (Cohen et al., 2006). A comparison study with school children found that those who participated in a year-long musical interaction group scored higher on an emotional empathy test than those who did not participate (Rabinowitch, Cross, & Burnard, 2013).

A key finding of this research is that all of the groups observed in this study experienced short-term benefits for socio-emotional wellbeing. These findings verify that the short-term effects of membership in both music and non-musical social groups are positive. Practically speaking, this research suggests that providing a broad range of social groups for older adults will result in improved emotion state and sense of social connection. Providing diverse socializing options that cater to a range of interests may be the best protective factor for aging well.

Contributorship

SM and NR both designed the study and collaborated on preparation of the manuscript and analysis of results. SM conducted the study and collected the data, and was the primary writer of the manuscript.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Peer review

Freya Bailes, University of Leeds, School of Music.

Karen Burland, University of Leeds, School of Music.

References

- Abell, R. V., Baird, A. D., & Chalmers, K. A. (2017). Group singing and health-related quality of life in Parkinson's disease. *Health Psychology, 36*, 55.
- Agahi, N., & Parker, M. G. (2008). Leisure activities and mortality: Does gender matter? *Journal of Aging and Health, 20*, 855–871.
- Allpress, R. B., Clift, S., & Legg, L. (2012). *Effects of group singing on psychological states and cortisol*. Paper presented at the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, Thessaloniki, Greece.
- Bailey, B. A., & Davidson, J. W. (2005). Effects of group singing and performance for marginalized and middle-class singers. *Psychology of Music, 33*, 269–303.
- Bailey, B. A., & Davidson, J. W. (2013). Emotional, social, and cognitive enrichment through participation in group singing: Interviews with members of a choir for homeless men. *The Phenomenon of Singing, 3*, 24–32.
- Baird, A., Abell, R., Haertsch, M., Bullott, N., Thompson, W. F., & Chalmers, K. (2015). *Group singing enhances positive affect in persons with Parkinson's Disease*. Paper presented at the The Art & Science of Music, Australian Music Psychology Society, Sydney, Australia.
- Bartel, C. A., & Saavedra, R. (2000). The collective construction of work group moods. *Administrative Science Quarterly, 45*, 197–231.
- Bartholomew, J. B., & Miller, B. M. (2002). Affective responses to an aerobic dance class: The impact of perceived performance. *Research Quarterly for Exercise and Sport, 73*, 301–309.
- Bennett, D. A., Schneider, J. A., Tang, Y., Arnold, S. E., & Wilson, R. S. (2006). The effect of social networks on the relation between Alzheimer's disease pathology and level of cognitive function in old people: A longitudinal cohort study. *The Lancet Neurology, 5*, 406–412.
- Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. *Trends in Cognitive Sciences, 17*, 179–193.
- Chin, T., & Rickard, N. (2012). The Music USE (MUSE) questionnaire: An instrument to measure engagement in music. *Music Perception, 29*, 429–446.
- Chirico, A., Serino, S., Cipresso, P., Gaggioli, A., & Riva, G. (2015). When music “flows”. State and trait in musical performance, composition and listening: A systematic review. *Frontiers in Psychology, 6*, 1–14.
- Clements-Cortés, A. (2015). Singing for health, connection and care. *Music and Medicine, 7*, 13–23.
- Clift, S., & Hancox, G. (2010). The significance of choral singing for sustaining psychological wellbeing: Findings from a survey of choristers in England, Australia and Germany. *Music Performance Research, 3*, 79–96.
- Clift, S., & Morrison, I. (2011). Group singing fosters mental health and wellbeing: Findings from the East Kent “singing for health” network project. *Mental Health and Social Inclusion, 15*, 88–97.
- Clift, S., Nicol, J., Raisbeck, M., Whitmore, C., & Morrison, I. (2010). Group singing, wellbeing and health: A systematic review. *UNESCO Observatory, 2*(1), 1–25.
- 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. *The Gerontologist, 46*, 726–734.
- Cornwell, B., Laumann, E. O., & Schumm, L. P. (2008). The social connectedness of older adults: A national profile. *American Sociological Review, 73*, 185–203.
- Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed.) (pp. 91–115). New York, NY: Guilford.
- Croom, A. M. (2015). Music practice and participation for psychological well-being: A review of how music influences positive emotion, engagement, relationships, meaning, and accomplishment. *Musicae Scientiae, 19*, 44–64.
- Cross, I. (2009). The evolutionary nature of musical meaning. *Musicae Scientiae, 13*, 179–200.
- Cross, I., Laurence, F., & Rabinowitch, T.-C. (2010). Empathy and creativity in group musical practices: Towards a concept of empathic creativity. In G. E. McPherson & G. F. Welch (Eds.), *The Oxford Handbook of Music Education* (pp. 1–21). Oxford: Oxford University Press.
- Csikszentmihalyi, M., & Larson, R. (2014). Validity and reliability of the experience-sampling method. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology* (pp. 35–54). Dordrecht: Springer.
- Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/Psychologie Canadienne, 49*, 182.
- Dingle, G. A., Brander, C., Ballantyne, J., & Baker, F. A. (2013). ‘To be heard’: The social and mental health benefits of choir singing for disadvantaged adults. *Psychology of Music, 41*, 405–421.
- Dingle, G. A., Williams, E., Jetten, J., & Welch, J. (2017). Choir singing and creative writing enhance emotion regulation in adults with chronic mental health conditions. *British Journal of Clinical Psychology, 56*, 443–457.
- Dubé, L., Chebat, J. C., & Morin, S. (1995). The effects of background music on consumers' desire to affiliate in buyer-seller interactions. *Psychology & Marketing, 12*, 305–319.
- Dunbar, R. I. M., Kaskatis, K., MacDonald, I., & Barra, V. (2012). Performance of music elevates pain threshold and positive affect: Implications for the evolutionary function of music. *Evolutionary Psychology, 10*, 688–702.

- Egermann, H., & McAdams, S. (2013). Empathy and emotional contagion as a link between recognized and felt emotions in music listening. *Music Perception: An Interdisciplinary Journal*, *31*, 139–156.
- Eisenberg, N., Eggum, N. D., & Di Giunta, L. (2010). Empathy-related responding: Associations with prosocial behavior, aggression, and intergroup relations. *Social Issues and Policy Review*, *4*, 143–180.
- Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, *101*, 91.
- Ekkekakis, P., Backhouse, S. H., Gray, C., & Lind, E. (2008). Walking is popular among adults but is it pleasant? A framework for clarifying the link between walking and affect as illustrated in two studies. *Psychology of Sport and Exercise*, *9*, 246–264.
- Fancourt, D., Williamson, A., Carvalho, L. A., Steptoe, A., Dow, R., & Lewis, I. (2016). Singing modulates mood, stress, cortisol, cytokine and neuropeptide activity in cancer patients and carers. *e-Cancer Medical Science*, *10*, 1–13.
- Fiori, K. L., Antonucci, T. C., & Cortina, K. S. (2006). Social network typologies and mental health among older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *61*, P25–P32.
- Fritz, B. S., & Avsec, A. (2007). The experience of flow and subjective well-being of music students. *Horizons of Psychology*, *16*, 5–17.
- Greenberg, D. M., Rentfrow, P. J., & Baron-Cohen, S. (2015). Can music increase empathy? Interpreting musical experience through the Empathizing–Systemizing (ES) theory: Implications for autism. *Empirical Musicology Review*, *10*, 80–95.
- Grewe, O., Kopiez, R., & Altenmüller, E. (2009). The chill parameter: Goose bumps and shivers as promising measures in emotion research. *Music Perception: An Interdisciplinary Journal*, *27*, 61–74.
- Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Listening to music as a re-creative process: Physiological, psychological, and psychoacoustical correlates of chills and strong emotions. *Music Perception*, *24*, 297–314.
- Groarke, J. M., & Hogan, M. J. (2015). Enhancing wellbeing: An emerging model of the adaptive functions of music listening. *Psychology of Music*, *44*, 769–791.
- Guhn, M., Hamm, A., & Zentner, M. (2007). Physiological and musico-acoustic correlates of the chill response. *Music Perception: An Interdisciplinary Journal*, *24*, 473–484.
- Hinshaw, T., Clift, S., Hulbert, S., & Camic, P. M. (2015). Group singing and young people's psychological well-being. *International Journal of Mental Health Promotion*, *17*(1), 46–63.
- Hirokawa, E. (2004). Effects of music listening and relaxation instructions on arousal changes and the working memory task in older adults. *Journal of Music Therapy*, *41*, 107–127.
- Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: a meta-analytic review. *Perspectives on Psychological Science*, *10*, 227–237.
- Horowitz, B. P., & Vanner, E. (2010). Relationships among active engagement in life activities and quality of life for assisted-living residents. *Journal of Housing for the Elderly*, *24*, 130–150.
- House, J. S., Landis, K. R., & Umberson, D. (1988). Social relationships and health. *Science*, *241*, 540–545.
- 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.
- Juslin, P. N., & Sloboda, J. (2010). *Handbook of music and emotion: Theory, research, applications*. Oxford, UK: Oxford University Press.
- Karageorghis, C. I., & Priest, D.-L. (2012). Music in the exercise domain: A review and synthesis (Part II). *International Review of Sport and Exercise Psychology*, *5*, 67–84.
- Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*, 354–364.
- Koelsch, S. (2013). From social contact to social cohesion—The 7 Cs. *Music and Medicine*, *5*, 204–209.
- Koys, D. J., & DeCotiis, T. A. (1991). Inductive measures of psychological climate. *Human Relations*, *44*, 265–285.
- Kreutz, G. (2014). Does singing facilitate social bonding? *Music and Medicine*, *6*, 51–60.
- Kreutz, G., Bongard, S., Rohrmann, S., Hodapp, V., & Grebe, D. (2004). Effects of choir singing or listening on secretory immunoglobulin A, cortisol, and emotional state. *Journal of Behavioral Medicine*, *27*, 623–635.
- Lang, P. J., & Bradley, M. M. (2010). Emotion and the motivational brain. *Biological Psychology*, *84*, 437–450.
- Lang, P. J., & Davis, M. (2006). Emotion, motivation, and the brain: Reflex foundations in animal and human research. *Progress in Brain Research*, *156*, 3–29.
- Laukka, P. (2007). Uses of music and psychological well-being among the elderly. *Journal of Happiness Studies*, *8*, 215–241.
- Lim, H. A. (2008). The effect of personality type and musical task on self-perceived arousal. *Journal of Music Therapy*, *45*, 147–164.
- Luanagh, C. Ó., & Lawlor, B. A. (2008). Loneliness and the health of older people. *International Journal of Geriatric Psychiatry*, *23*, 1213–1221.
- Luo, Y., Hawkey, L. C., Waite, L. J., & Cacioppo, J. T. (2012). Loneliness, health, and mortality in old age: A national longitudinal study. *Social Science & Medicine*, *74*, 907–914.
- MacDonald, R. A. (2013). Music, health, and well-being: A review. *International Journal of Qualitative Studies on Health and Well-being*, *8*, 20635. doi: 10.3402/qhw.v8i0.20635
- Maurly, S., & Rickard, N. (2016). Wellbeing in the classroom: How an evolutionary perspective on human musicality can inform music education. *Australian Journal of Music Education*, *50*, 3–15.
- Menec, V. H. (2003). The relation between everyday activities and successful aging: A 6-year longitudinal study. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *58*, S74–S82.

- Miu, A. C., & Balteş, F. R. (2012). Empathy manipulation impacts music-induced emotions: A psychophysiological study on opera. *PLoS One*, *7*, e30618.
- Nakamura, J., & Csikszentmihalyi, M. (2014). The concept of flow. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology* (pp. 239–263). Dordrecht: Springer.
- Niedenthal, P. M., & Brauer, M. (2012). Social functionality of human emotion. *Annual Review of Psychology*, *63*, 259–285.
- Ortiz, J. M. (1997). *The Tao of music: Using music to change your life*. Dublin: Newleaf.
- Páez, D., Rimé, B., Basabe, N., Włodarczyk, A., & Zumeta, L. (2015). Psychosocial effects of perceived emotional synchrony in collective gatherings. *Journal of Personality and Social Psychology*, *108*, 711.
- Pearce, E., Launay, J., & Dunbar, R. I. M. (2015). The ice-breaker effect: Singing mediates fast social bonding. *Royal Society Open Science*, *2*(10), 1–9. doi:10.1098/rsos.150221
- Pearce, E., Launay, J., MacCarron, P., & Dunbar, R. I. (2017). Tuning in to others: Exploring relational and collective bonding in singing and non-singing groups over time. *Psychology of Music*, *45*, 496–512.
- Potočnik, K., & Sonnentag, S. (2013). A longitudinal study of well-being in older workers and retirees: The role of engaging in different types of activities. *Journal of Occupational and Organizational Psychology*, *86*, 497–521.
- Rabinowitch, T.-C., Cross, I., & Burnard, P. (2013). Long-term musical group interaction has a positive influence on empathy in children. *Psychology of Music*, *41*, 484–498.
- Reddish, P., Bulbulia, J., & Fischer, R. (2014). Does synchrony promote generalized prosociality? *Religion, Brain & Behavior*, *4*, 3–19.
- Rickard, N. S. (2004). Intense emotional responses to music: A test of the physiological arousal hypothesis. *Psychology of Music*, *32*, 371–388.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68.
- Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*, *14*, 257–262.
- Salimpoor, V. N., Benovoy, M., Longo, G., Cooperstock, J. R., & Zatorre, R. J. (2009). The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS One*, *4*, e7487.
- Sanal, A. M., & Gorsev, S. (2014). Psychological and physiological effects of singing in a choir. *Psychology of Music*, *42*, 420–429.
- Sandgren, M. (2009). *Evidence for strong immediate well-being effects of choral singing—With more enjoyment for women than for men*. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & P.-S. Eerola (Eds.), *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009)*, Jyväskylä: University of Jyväskylä.
- Schellenberg, E. G., Nakata, T., Hunter, P. G., & Tamoto, S. (2007). Exposure to music and cognitive performance: Tests of children and adults. *Psychology of Music*, *35*, 5–19.
- Schulkin, J., & Raglan, G. B. (2014). The evolution of music and human social capability. *Frontiers in Neuroscience*, *8*(292), 1–13.
- Seeman, T. E. (2000). Health promoting effects of friends and family on health outcomes in older adults. *American Journal of Health Promotion*, *14*, 362–370.
- Seeman, T. E., Lusignolo, T. M., Albert, M., & Berkman, L. (2001). Social relationships, social support, and patterns of cognitive aging in healthy, high-functioning older adults: MacArthur studies of successful aging. *Health Psychology*, *20*, 243.
- Sinnamon, S., Moran, A., & O’Connell, M. (2012). Flow among musicians: Measuring peak experiences of student performers. *Journal of Research in Music Education*, *60*, 6–25.
- Steptoe, A., Shankar, A., Demakakos, P., & Wardle, J. (2013). Social isolation, loneliness, and all-cause mortality in older men and women. *Proceedings of the National Academy of Sciences*, *110*, 5797–5801.
- Stewart, N. A. J., & Lonsdale, A. J. (2016). It’s better together: The psychological benefits of singing in a choir. *Psychology of Music*, *44*(6), 1240–1254.
- Stratton, V. N., & Zalanowski, A. H. (1984a). The effect of background music on verbal interaction in groups. *Journal of Music Therapy*, *21*, 16–21.
- Stratton, V. N., & Zalanowski, A. H. (1984b). The relationship between music, degree of liking and self-reported relaxation. *Journal of Music Therapy*, *21*, 184–192.
- Sze, J. A., Gyurak, A., Goodkind, M. S., & Levenson, R. W. (2012). Greater emotional empathy and prosocial behavior in late life. *Emotion*, *12*, 1129.
- Telle, N.-T., & Pfister, H.-R. (2015). Positive empathy and prosocial behavior: A neglected link. *Emotion Review*, *8*(2), 154–163.
- Thayer, R. E. (1978). Factor analytic and reliability studies on the activation-deactivation adjective check list. *Psychological Reports*, *42*, 747–756.
- Thayer, R. E. (1986). Activation-deactivation adjective check list: Current overview and structural analysis. *Psychological Reports*, *58*, 607–614.
- Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science*, *12*, 248–251.
- Trainor, L. J., & Cirelli, L. (2015). Rhythm and interpersonal synchrony in early social development. *Annals of the New York Academy of Sciences*, *1337*, 45–52.
- Unger, J. B., McAvay, G., Bruce, M. L., Berkman, L., & Seeman, T. (1999). Variation in the impact of social network characteristics on physical functioning in elderly persons: MacArthur Studies of Successful Aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *54*, S245–S251.
- Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion*, *11*, 262.

- Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of Experimental Social Psychology, 46*, 693–695.
- van der Zwaag, M. D., Westerink, J. H., & van den Broek, E. L. (2011). Emotional and psychophysiological responses to tempo, mode, and percussiveness. *Musicae Scientiae, 15*, 250–269.
- von Lob, G., Camic, P., & Clift, S. (2010). The use of singing in a group as a response to adverse life events. *International Journal of Mental Health Promotion, 12*, 45–53.
- Walker, C. J. (2010). Experiencing flow: Is doing it together better than doing it alone? *The Journal of Positive Psychology, 5*, 3–11.
- Watson, D., Clark, I. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology, 47*, 1063–1070.
- Welch, G. F., Himonides, E., Saunders, J., Papageorgi, I., & Sarazin, M. (2014). Singing and social inclusion. *Frontiers in Psychology, 5*(803), 1–12.
- Ziv, N., Granot, A., Hai, S., Dassa, A., & Haimov, I. (2007). The effect of background stimulative music on behavior in Alzheimer's patients. *Journal of Music Therapy, 44*, 329–343.