Group music performance causes elevated pain thresholds and social bonding in small and large groups of singers

Daniel Weinstein[a], **Jacques Launay**[b], **Eiluned Pearce**[b], **Robin I. M. Dunbar**[b], and **Lauren Stewart**[a]

Jacques Launay: jacques.launay@psy.ox.ac.uk; Eiluned Pearce: eiluned.pearce@psy.ox.ac.uk; Robin I. M. Dunbar: robin.dunbar@psy.ox.ac.uk; Lauren Stewart: l.stewart@gold.ac.uk

[a]Goldsmiths, University of London, Lewisham Way, New Cross, London SE14 6NW, United Kingdom, +44 20 7919 7171

[b]University of Oxford, Department of Experimental Psychology, Tinbergen Building, Oxford, OX1 3UD, United Kingdom, +44 1865 271367

Abstract

Over our evolutionary history, humans have faced the problem of how to create and maintain social bonds in progressively larger groups compared to those of our primate ancestors. Evidence from historical and anthropological records suggests that group music-making might act as a mechanism by which this large-scale social bonding could occur. While previous research has shown effects of music making on social bonds in small group contexts, the question of whether this effect 'scales up' to larger groups is particularly important when considering the potential role of music for large-scale social bonding. The current study recruited individuals from a community choir that met in both small (n = 20 – 80) and large (a 'megachoir' combining individuals from the smaller subchoirs n = 232) group contexts. Participants gave self-report measures (via a survey) of social bonding and had pain threshold measurements taken (as a proxy for endorphin release) before and after 90 minutes of singing. Results showed that feelings of inclusion, connectivity, positive affect, and measures of endorphin release all increased across singing rehearsals and that the influence of group singing was comparable for pain thresholds in the large versus small group context. Levels of social closeness were found to be greater at pre- and post-levels for the small choir condition. However, the large choir condition experienced a greater change in social closeness as compared to the small condition. The finding that singing together fosters social closeness – even in large contexts where individuals are not known to each other – is consistent with evolutionary accounts that emphasize the role of music in social bonding, particularly in the context of creating larger cohesive groups than other primates are able to manage.

Keywords

music; endorphins; social bonding; evolution; singing

Daniel Weinstein (corresponding author): ps301dw@gold.ac.uk, (1) 845 636 9482.

1 Introduction

Membership of a social group has been imperative for survival throughout human evolution and continues to exert a substantial influence on individual health and well-being today (e.g. Umberson and Montez, 2010). In fact, the beneficial effects of a supportive social network on life expectancy is of the same magnitude as well-known health-promotion factors such as giving up smoking and doing exercise (Holt-Lunstad et al., 2010). Sustaining social relationships with members of our network and larger cultural groups is therefore vitally important, yet the mechanisms by which we do this in large social groups remain only poorly understood.

In non-human primates, social ties are created and sustained through intense one-on-one grooming, which stimulates the release of endorphins in the brain and promotes emotional closeness between the individuals involved (Keverne et al. 1989; Machin & Dunbar 2011). However, given that there are other essential activities (such as feeding) to fit into a day, the amount of time it takes to groom limits the number of such relationships that can be maintained (Dunbar, 1991; Lehmann, Korstjens, & Dunbar, 2007). Since humans live in much larger groups than other primates, sustaining greater numbers of relationships within a finite time budget requires mechanisms that facilitate bonding with several individuals simultaneously (e.g. Aiello & Dunbar, 1993; Dunbar, 2003; Dunbar, 2008). Our hunter-gatherer ancestors needed to exploit behaviours that have equivalent social bonding effects to physical grooming, but which can be performed with several other individuals at the same time, rather than being limited to one-on-one interactions.

Human hunter-gatherers spend the majority of their time living in residential bands of up to 50 individuals that allow cooperative foraging and collective childcare, but these bands seasonally aggregate to form 'mega bands' numbering several hundred individuals (Binford, 2001; Dunbar, 1998; Hill & Dunbar, 2003; Zhou, Sornette, Hill, & Dunbar, 2005; Hamilton et al., 2007). These periodic gatherings have been interpreted as opportunities to create and maintain social ties with members of other bands that habitually range in distant areas and are infrequently encountered (Pearce, 2014; Pearce, Shuttleworth, Grove, & Layton, 2014). Such long-distance ties provide a safety net against local resource failure by bringing information about resources in other areas and also allow access to mates outside one's own band (Whallon, 2006; Lehmann, Lee, & Dunbar 2014). The activities performed in these short-term mega-bands include group rituals which often include singing and dancing (e.g. Guemple, 1971), and these behaviours have commonly been interpreted as a means to create and maintain affiliative ties that by-passes the need for all individuals to personally interact with one another.

Given the widespread occurrence of musical behaviour both cross-culturally and historically and the fact that the vast majority of individuals have some musical ability (Sloboda, Davidson, & Howe, 1994; Kalender, 2013; Launay, Grube & Stewart, 2014), it has been argued that music is an evolutionary adaptation (e.g. Dunbar, Kaskatis, Macdonald, & Barra, 2012; Huron, 2012; Tarr, Launay, Dunbar, 2014). Previous research has demonstrated that music can provide a sense of group identity (Laiho, 2004; Miranda & Gaudreau, 2011) and foster prosocial behaviour (Bakagiannis & Tarrant, 2006) and it is likely that the particular

combination of activities involved in music-making encourage a sense of social closeness within groups. The initial shared motivation required to produce music (e.g. Reddish, Fischer & Bulbulia, 2013), the shared attention involved (e.g. Shteynberg, Hirsch, Galinsky & Knight, 2014; Wolf, Launay & Dunbar, 2015), the act of predicting movements of another person (Sebanz & Knoblich, 2009), physical coordination itself (e.g. Hove & Risen, 2009) and a sense of shared success (e.g. Launay, Dean & Bailes, 2013) are all in themselves likely to lead to greater affinity with other group members. Music-making can thus be construed as a technology which brought all these elements together, and conferred a selective advantage for social bonding behaviour.

Psychological research into this social bonding effect is supported by evidence that music-making results in endorphin release, akin to that seen in social bonding induced by grooming in other primate species (Dunbar et al., 2012b; Fabre-Nys, Meller, & Keverne, 1982; Keverne et al., 1989; Machin & Dunbar, 2011; Martel, Nevison, Simpson, & Keverne, 1995; Meller, Keverne, & Herbert, 1980). In non-human primate species, levels of endorphins are elevated after grooming (Keverne et al., 1989; Martel et al. 1995), and blocking endorphin receptors leads to increased grooming rates (Meller et al., 1980; Fabre-Nys, 1982). Since direct measures of endorphins require logistically and ethically problematic procedures (PET scanning or lumbar puncture) and given the known role of endorphins as an endogenous analgesic (Zubieta et al., 2001; Zubieta et al., 2003; Hsu et al., 2013), pain threshold measures are commonly used as a proxy in humans (e.g. Cohen et al., 2010; Dunbar et al., 2012a). With this measure, it has been shown that synchronised activities involving muscular exertion lead to a larger increase in pain thresholds (interpreted as greater endorphin release) compared with unsynchronised movements (Sullivan, Rickers, & Gammage, 2014). In addition, increased pain thresholds are found when similar movements are performed as a coordinated group rather than alone (Cohen et al., 2010). Moreover, synchrony seems to have the same effect regardless of whether the activity is performed in the presence of strangers or known others (Sullivan & Rickers, 2013). Similarly, both laughter and active music-making (singing and drumming), which often occur in coordination with others and require muscular exertion to produce, have been shown to lead to an increase in pain thresholds compared with control activities (Dunbar et al, 2012a; Dunbar et al, 2012b). Overall, coordinated music-making, such as singing, is a likely candidate for a behaviour that would allow multiple individuals to be 'groomed' at the same time (and experience endorphin release), thus allowing larger groups to be maintained.

In this paper, we address the possibility that music-making may represent a form of social interaction that allows cohesive groups to expand beyond the limit imposed when social connectivity has to rely on a series of one-to-one grooming interactions. This cannot be taken for granted, since previous work demonstrates that there are limits on the number of people who can engage simultaneously in other bonding behaviours, such as laughter (Dezecache & Dunbar, 2012).

In order to investigate this, we worked with members of 'Popchoir' (http://www.popchoir.com) which comprises 10 subchoirs that meet in their local context on a weekly basis, but gather together once a year to perform as a megachoir comprising members across different subchoirs. Critically, the subchoirs all learn the same repertoire,

allowing them to perform en masse when they gather as the 'megachoir.' The group sizes involved mirror the sizes of hunter-gatherer bands and periodic aggregations into mega-bands (Binford, 2001), allowing us to investigate whether singing is similarly able to bond smaller groups of familiar individuals (the equivalent size of a band: 20 – 80 members) and large groups of familiar and unfamiliar individuals (equivalent in size to periodic aggregations of bands numbering several hundred individuals). This allows us to test whether singing can bond social groups beyond the size observed in non-human primates: that is, to assess whether the social bonding effects of music-making apply to large as well as small groups. A secondary interest was the extent to which bonding in the megachoir setting even extends between individuals who are not known to one another. For instance, it is possible that in the megachoir situation, individuals experience a sense of community but that this only leads to greater cohesion with the members of their existing local group ('subchoir') rather than the entire 'megachoir' – potentially an in-group/out-group division.

To examine the effect of group size on the bonding effects of singing, we tested the following hypotheses: (1) participants will report an increase in social bonding as a consequence of group singing, as well as an increase in positive affect and decrease in negative affect, and an increase in pain thresholds (replicating previous results as a check of validity; Dunbar et al., 2012b); (2) singing in a large group condition will replicate the changes seen in the small group condition, indicating that the social bonding effects of singing "scale up" and the absence of a bonding limit below several hundred group members. We also tested whether people report more social bonding towards members of their own familiar subchoirs compared with the whole megachoir when they all rehearse together. While we had no specific hypotheses regarding this latter comparison, the results are relevant to whether singing is effective at bonding very large groups of people who do not already know each other, or whether it is more effective at enhancing existing social bonds within a smaller, more intimate group of known individuals.

2 Methods

2.1 Participants

Participants included members of a non-professional choir known as Popchoir, which is made up of 10 sub-choirs around the greater London area, each of which contains approximately 20 to 80 individuals. All members of each sampled choir were asked to fill out self-report measures of affect and social connectedness, and some members also volunteered for pain threshold tests (referred to as 'pressure cuff tests' to participants). Six of the 10 subchoirs were sampled during individual rehearsals for both self-report data and pain threshold tests and their responses were amalgamated; this constituted the "small" group condition. The singers of Popchoir each attend only a single subchoir. Additionally, these 10 sub-choirs rehearse together a few times per year as one larger, combined choir. This combined rehearsal comprised of 232 singers and constituted the "large" group condition during which data was collected in the same manner as the small group condition.

In the small group condition, 133 participants completed the self-report measures across the six choir groups (118 female, 15 male; age: range 21 – 68, $M = 42.5$, $SD = 12.4$) and 45 participants contributed pain threshold data (38 female, 7 male; age: range 24 – 66, $M =$

44.82, $SD = 11.88$). In the large group condition, 80 participants completed the self-report measures (71 female, 9 male; age: range 21-75, $M = 43.2$, $SD = 12.2$) and 62 gave pain threshold data (53 female, 9 male; age: range 28-75, $M = 48.05$, $SD = 11.78$)[2]. Participants who failed to complete the self-report measures correctly were excluded. Following Dunbar et al. (2012b), participants who were pregnant, had smoked or consumed alcohol in the last two hours were excluded from the pain threshold data. While Cohen et al. (2010) and Dunbar et al. (2012b) used time as a measurement for pain thresholds with a 180 second limit, we measured pressure applied with a maximum possible pain threshold measure of 300mmHg (using standard blood pressure cuffs). Three participants were excluded because they had an initial measurement of 300 mmHg and were therefore at ceiling level on the baseline measure. Participants were not made aware of the nature of our investigation.

2.2 Materials and Design

In addition to background information (age, gender, length of time in choir, familiarity with the choir on a Likert Scale (1 – 7)), the self-report survey consisted of three elements that were completed both before and after the rehearsal: positive and negative affect (short form of the PANAS; see Mackinnon, 1998), IOS scale (Inclusion of Other in Self scale; see Aron, Aron, & Smollan, 1992), and a measure of social connectedness ("connectivity"). The PANAS contains ten different words that reflect either positive or negative affect (five of each) with a 1 – 5 Likert Scale of how the participant is feeling at that particular moment, with 1 representing little to none of that emotion and 5 representing an extreme amount of that emotion (Mackinnon, 1998). The IOS scale was presented in the form of Venn Diagrams that pictorially represent inclusion of oneself in a group scored as a Likert scale ranging from 1 – 7; 1 represents extreme exclusion and 7 represents extreme inclusion (Aron et al., 1992). The social connectedness question was presented as a Likert scale ranging from 1 (not connected at all) to 7 (extremely connected). Participants circled the number that best matched their current state for each item. Participants in the large group condition were additionally asked about their social relationships to the choir with which they normally sing (subchoir), as well as their relationships to the megachoir as a whole. Thus, feelings of inclusion and connectivity were collected from participants towards their own choir when in the small group condition and towards both their own choir and towards the megachoir as a whole when in the large group condition. All other items were measured in only the context of the small or large group condition.

The dependent variables – self-report measures and pain thresholds – were taken from participants before and after 90 minutes of singing, creating a within-subjects comparison. The independent variable was group size, and consisted of the small group condition (less than 80 individuals) and large group condition (232 individuals).

2.3 Procedure

Information sheets, consent forms, and the self-report forms were distributed before the start of each rehearsal. Choir members completed these forms before rehearsal began. Volunteers

---

[2]Given that all participant data were anonymous, it is not possible to say how many participants contributed only self-report data, only pain threshold data, or both.

for the pain threshold tests and were taken to a quiet area away from the other choir members. In the small group condition, between 5 and 10 participants were tested in each of the 6 choir groups. A standard issue mercurial sphygmomanometer (blood pressure cuff) was attached just above the elbow of the participant's non-dominant arm and the experimenter stated, "I will inflate the cuff slowly. Please let me know when it is very uncomfortable by saying 'now.'" The experimenter then began inflating the cuff steadily in 10 mmHg increments until the participant stated that they were very uncomfortable. The experimenter then recorded the pressure, rounding to the nearest 5 mmHg. Pressure applied to participants could not exceed 300 mmHg.

After the initial measures had been taken, the rehearsals proceeded as normal. Popchoir's repertoire consists mainly of music in the pop genre, for example, "Halfway There" by Bon Jovi, "Spice Up Your Life" by the Spice Girls, "Skyscraper" by Sam Bailey, and "The Circle of Life" by Elton John. After approximately 90 minutes of rehearsing, participants were asked to complete the second part of the self-report questionnaire, and participants who had previously contributed pain threshold data were asked to return for their post-rehearsal measurements, following the same procedure as the initial measurements.

2.4 Variable calculation and analysis

Positive affect items from the PANAS were added together, and likewise for the negative affect items, creating separate positive and negative affect scores per individual. Difference scores were calculated for all self-report measures (PANAS, IOS and connectivity) and pain threshold data by subtracting the measures taken before singing from those taken after the singing. The IOS, connectivity and PANAS scores were not normally distributed (Kolmogorov-Smirnov tests, $p < .05$), so non-parametric Mann Whitney and Wilcoxon tests were used for their analysis[3]. Pain threshold data did not deviate from a normal distribution (Kolmogorov-Smirnov test, $p > .05$) and so a parametric test (two-way mixed ANOVA) was used to analyse these data. Intraclass correlations within singing classes were not significant for all of the difference scores except positive PANAS, so in this case, multilevel linear modelling with class as level was used to analyse the data.

3 Results

3.1 Small choir (20-80 individuals) versus large choir (232 individuals)

Descriptive statistics for IOS, connectivity, and positive and negative affect are given in Table 1. To determine if values of these self-report measures changed as a result of singing, Wilcoxon One Sample Signed Rank tests using a test value of 0 were carried out on all four measures in both the small and large conditions. In the small condition, results were significantly different from zero for the IOS difference, $Mdn = 1$, $Z = -7.6$, $p < .01$, $r = .47$, as well as the connectivity difference, $Mdn = 1$, $Z = -7.1$, $p < .001$, $r = .44$. Change in positive affect was also significantly different from zero, $Mdn = 3$, $Z = -7.9$, $p < .001$, $r = .49$, as was negative affect, $Mdn = 0$, $Z = -6.0$, $p < .001$, $r = .37$. Similarly, in the large condition, change in IOS was significantly different from zero, $Mdn = 1$, $Z = -6.0$, $p < .001$,

---

[3] ANOVAs yielded similar results.

$r = .48$, as was connectivity, $Mdn = 2$, $Z = -7.4$, $p < .001$, $r = .59$. Finally, in the large group condition, change in positive affect was also significantly different from zero, $Mdn = 2$, $Z = -5.3$, $p < .001$, $r = .42$, as well as change in negative affect, $Mdn = 0$, $Z = -3.6$, $p < .001$, $r = .28$.

Changes in IOS scores were significantly larger in the large group condition ($Mdn = 1$, min = -3, max = 4) compared to the small group condition ($Mdn = 1$, min = -3, max = 5), $U = 3988$, $Z = -3.2$, $p = .001$ (two-tailed), $r = .22$, as shown in Figure 1. Similarly, changes in connectivity in the large group condition ($Mdn = 1$, min = -1, max = 4) were significantly larger than those in the small group condition ($Mdn = 2$, min = -3, max = 3), $U = 2386$, $Z = -7.0$, $p < .001$, $r = .48$ (Figure 1). Changes in positive affect were significantly greater in the small group condition ($Mdn = 3$, min = -6, max = 11) compared with the large group condition using Mann Whitney tests ($Mdn = 2$, min = -5, max = 13), $U = 3965$, $Z = -3.1$, $p = .002$, $r = .21$. However, change in positive affect showed a significant intraclass correlation within the different choir groups in the small group condition ($r = 0.07$, 95% CI [0.002, 0.43]) suggesting participants in the same subchoir were more likely to experience similar change in positive affect. We therefore used a hierarchical linear model to include class as a level and including this suggested that the change in positive affect was only marginally higher in the small choir condition compared with the large choir condition ($F(1,150) = 3.26$, $p = 0.07$). There was no significant difference in the change in negative affect between conditions, $U = 4603.5$, $Z = -1.9$, $p = .061$ (Figure 2).

While the relative differences (pre vs. post), as function of small versus large group condition were of primary interest, we were also motivated to understand where these differences had emerged. To address this question, we compared raw scores of IOS and connectivity both before and after singing using a Mann-Whitney test. These demonstrated that at baseline levels, there were significantly higher ratings for IOS in the small choir condition ($Mdn = 4$, min = 1, max = 7) compared with the large choir condition ($Mdn = 2$, min = 1, max = 7), $U = 3392.5$, $Z = -4.54$ $p < .001$, $r = .31$. Similarly, connectivity was also significantly greater in the small condition ($Mdn = 4$, min = 1, max = 7) than in the large condition (Mdn = 4, min = 1, max = 7), $U = 3574$, $Z = -4.11$, $p < .001$, $r = .28$. After singing, there were also significantly higher ratings of IOS in the small condition ($Mdn = 4$, min = 1, max = 7), compared with the large choir condition ($Mdn = 4$, min = 1, max = 7), $U = 4263.5$, $Z = -2.47$, $p = .014$, $r = .17$. Connectivity followed the same pattern; measures were greater in the small condition ($Mdn = 5$, min = 1, max = 7) when compared to the large condition ($Mdn = 5$, min = 1, max = 5), $U = 4223$, $Z = -2.60$, $p = .009$, $r = .18$, indicating that although ratings increased more in the large choir condition, they did not reach the same level as ratings in the small choir condition after singing. Table 1 shows the descriptive statistics for all self-report measures.

For pain thresholds, a significant main effect of time point was found, $F(1, 105) = 14.2$, $p < .001$, $G\eta^2 = 0.01$, but no effect of condition was found, $F(1, 105) = 1.5$, $p = .23$, nor was there a significant interaction, $F(1, 105) = 0.37$, $p = .54$. Thus, pain tolerance levels increased after singing, regardless of group size condition (Figure 3).

3.2 Subchoir versus megachoir choir

Given that participants in the large condition answered questions regarding their primary choir (subchoir measure) as well as the combined choir that attended the rehearsal (megachoir measure), we also compared these variables, creating a within-subjects analysis. A Wilcoxon matched pairs test was used to compare these subchoir and megachoir measurements. The megachoir (*Mdn* = 1, min = -3, max = 4) showed a significantly greater change in IOS when compared to the subchoir (*Mdn* = 0, min = -4, max = 3), *Z* = -5.0, *p* < .001, *r* = .40. Similarly, the change in connectivity for the megachoir (*Mdn* = 2, min = -1, max = 4) was significantly greater than that of the subchoir (*Mdn* = 0, min = -3, max = 2), *Z* = -6.9, *p* < .001, *r* = .55 (Figure 4).

Again, we examined pre-measurements as well as post-measurements using raw scores to further examine these differences with a Mann-Whitney test. At baseline levels, IOS for the subchoir was significantly greater (*Mdn* = 4, min = 1, max = 7) than in the megachoir (*Mdn* = 2, min = 1, max = 7), *U* = 1702.5, *Z* = -5.22, *p* < .001, *r* = .41. Connectivity showed the same result: the subchoir showed a significantly greater starting point (*Mdn* = 4, min = 1, max = 7) than the megachoir (*Mdn* = 4, min = 1, max = 7), *U* = 1960, *Z* = -4.34, *p* < .001, *r* = .34. However, these differences disappeared by the end of the rehearsal. Analysis showed that IOS post-measurements in the subchoir (*Mdn* = 4, min = 1, max = 7) were not significantly different from post-measurements in the megachoir (*Mdn* = 4, min = 1, max = 7), *U* = 2725.5, *Z* = -1.65, *p* = .10. Connectivity difference between the subchoir (*Mdn* = 5, min = 1, max = 7) was also non-significant from the megachoir (*Mdn* = 5, min = 1, max = 7), *U* = 2919, *Z* = -.99, *p* = .321, meaning that these two groups had post-measurements at approximately the same level.

3.3 Relationships between dependent variables

Finally, we examined the association between positive affect on both bonding measures by running a Spearman correlation. For the small condition, positive affect was significantly correlated with both IOS (r_s = .231, *p* = .007) and connectivity (r_s = .276, *p* = .001). In the large condition, connectivity (r_s = .294, *p* = .008), but not IOS (r_s = .191, *p* = .09), had a significant positive correlation with positive affect.

To further investigate whether positive affect was a predictor of bonding, we tested linear models for both bonding measures. For change in connectivity, a model including change in positive affect and condition was significant (*p* < 0.001, adj. r^2 = 0.11) with both change in positive affect (*b* = 0.10, *t*(2,210) = 4.79, *p* < 0.001), and condition acting as significant predictors (*b* = 0.45, *t*(2,210) = 3.261, *p* = 0.0382). Similarly, a model including both predictors was significant for change in IOS (*p* < 0.001, adj. r^2 = 0.08; change in positive affect: *b* = 0.09, *t*(2,210) = 3.72, *p* < 0.001; condition: *b* = 0.5, *t*(2,210) = 3.215, *p* = 0.002).

4 Discussion

We set out to test whether music might act as a mechanism by which large-scale social bonding can occur. Taking advantage of an existing choir structure ('Popchoir'), whereby small groups periodically aggregate to sing as a megachoir, we compared effects of music

making on self-reported social bonding measures, as well as pain thresholds (as a proxy for endorphin release). Because we are unaware of any existing non-music based groups that periodically aggregate into a 'mega-group' (as for Popchoir), the primary focus of the present study was to specifically examine the effect of music making in small versus large group contexts rather than to examine the music-specificity of the effect. We hypothesized that (1) participants would report an increase in social bonding following group singing, as well as an increase in positive mood, decrease in negative mood, and an increase in pain thresholds and (2) singing in a large group condition would demonstrate comparable changes in these measures compared with a small group condition.

Overall, we replicate previous findings that demonstrate that singing leads to increases in positivity, social bonding, and pain thresholds (e.g., Dunbar et al. 2012b; Getz, Chamorro-Premuzic, Roy, & Devroop, 2012; Kawakami, Furukawa, Katahira, & Okanoya, 2013; Kirschner, & Tomasello, 2010). In addition, the subjectively reported social bonding effects of singing were significantly greater in a large group (of 232 people) than in smaller groups (20-80 people) and the change in pain threshold was comparable across the small and large group conditions. Although initial analysis suggested that increases in positive affect were greater in the small group condition, these results were not robust once intraclass correlation was taken into account.

Pre-rehearsal levels of social closeness significantly differed between the conditions. The small condition and subchoirs began with a higher baseline level of bonding compared to that of the large choir condition. This is what one may expect since the small choirs meet once a week, whereas the megachoir meets infrequently (once or twice a year). Importantly, we show that even after only a single session of singing, a large group of unfamiliar individuals can become bonded to the same level as those who are familiar to each other within that group.

The fact that the large condition experienced a greater change in social bonding levels supports the hypothesis that music-making is a particularly powerful means by which to encourage bonding in large groups of people who are relatively unfamiliar with one another, and suggests that any limit on the group size that can be bonded in this manner lies beyond several hundred people. In our evolutionary history, the discovery of such a technology may have led to better intergroup relationships between neighbouring bands, and so facilitated the evolution of large scale fission-fusion sociality (Dunbar, 2003; Dunbar, 2012; Pearce et al., 2014). Our results suggest that communal singing can cause a significant increase in social closeness of large groups of unfamiliar individuals (c.f. Pearce, Launay & Dunbar, in press). In other words, communal singing may bypass the need for personal knowledge about other individuals that more intimate relationships require.

The effect of group singing on pain thresholds were found to be comparable, regardless of group size, suggesting that the physiological reaction (pain thresholds) to a period of singing is the same regardless of context, even if the psychological experience is different (self-reported bonding measures). The change in pain thresholds identified here mirror those previously associated with laughter (Dunbar et al., 2012a), rowing in synchrony (Cohen et al., 2010), and active musical engagement (Dunbar et al., 2012b) in social settings. These

analgesic effects have been interpreted as relating to the release of endorphins, thought to be important in mediating primate social bonding (Curley & Keverne, 2005; Depue & Morrone-Strupinsky, 2005; Dunbar, 2010; Machin & Dunbar, 2011; Taylor et al., 2013). Several experiments have demonstrated that behavioural co-ordination can increase cooperative behaviour (e.g. Cirelli, Einarson & Trainor, 2014; Launay, Dean & Bailes, 2014; Valdesolo & DeSteno, 2009; Wiltermuth & Heath, 2009), and this effect may well be mediated by the release of endorphins as well as other contributing psychological factors.

IOS and connectivity were correlated with positive affect in the small condition, but only IOS was correlated in the large condition. Since Dunbar et al. (2012a) previously demonstrated that music-making is associated with an increase in pain threshold independently of affect, these correlations between the bonding and affect measures might suggest two routes to social bonding: one via the endogenous opioid system and the second via increases in positive affect. For example, the stronger correlations between both social bonding measures and positive affect in the small group condition alone might demonstrate a dissociation between affect and feelings of social cohesion, with affect playing less of a role in social bonding in larger groups. With the present study design, however, we cannot determine whether differences between the small and large group conditions are due to the size of the large group per se or to the fact that many of the singers did not know one another. To reliably differentiate between these two possibilities, we would need to compare singing in large groups that consisted entirely of either familiar individuals or strangers but to do this in an ecologically valid context, as in the present study, will be challenging.

While naturalistic studies inevitably suffer from a lack of control over extraneous variables, they have the merit of providing more valid insights into the real-world function of singing than do laboratory paradigms. Using the current experimental design, we directly compared different sized singing groups but did not contrast these activities with a control activity, given that the periodic aggregation that we saw with Popchoir is not, to our knowledge, mirrored by any other community activity, and because previous research has already shown that musical activities can lead to more social bonding than control activities (e.g., Dunbar et al., 2012b). Further experimental work will be needed to determine the exact causal relationships between the self-report data and pain thresholds as a consequence of different social bonding activities.

Overall, we show that group singing can result in an increase in both positive affect, feelings of inclusion and social connection, and an increase in pain thresholds. More importantly, we show that the social bonding effects of singing are actually more substantial in larger group settings compared to smaller, more familiar groups. Though social closeness in the large condition did not meet the level found in the small condition, we may expect the relationships between individuals to grow over a longer period of time than a single rehearsal. Altogether, the results here are consistent with the hypothesis that music can act as a mechanism for the facilitation and maintenance of social connections between group members, including large groups made up of unfamiliar individuals. Moreover, any threshold on this effect seems to lie beyond group sizes of ~200 individuals. This supports the notion that diverse cultural phenomena such as national anthems, religious music, team chants, or military marching bands are behaviours that promote social bonding in large

groups of individuals who do not necessarily know each other personally. Such behaviours may have played a crucial role in human evolution by allowing us to increase community size significantly beyond those found in other primate species.

Acknowledgments

References

Aiello LC, Dunbar RIM. Neocortex size, group size, & the evolution of language. Current Anthropology. 1993; 34:184–193.

Aron A, Aron EN, Smollan D. Inclusion of Other in the Self Scale and the structure of interpersonal closeness. Journal of Personality and Social Psychology. 1992; 63(4):596–612.

Bakagiannis S, Tarrant M. Can music bring people together? Effects of shared musical preference on intergroup bias in adolescence. Scandinavian Journal of Psychology. 2006; 47(2):129–36. [PubMed: 16542355]

Binford, LR. Constructing Frames of Reference: An Analytical Method for Archaeological Theory Building Using Ethnographic and Environmental Data Sets. University of California Press; 2001.

Cirelli LK, Einarson KM, Trainor LJ. Interpersonal synchrony increases prosocial behavior in infants. Developmental Science. 2014

Cohen EEA, Ejsmond-frey R, Knight N, Dunbar RIM. Rowers ' high : behavioural synchrony is correlated with elevated pain thresholds. Biology Letters. 2010; 6(10):106–8. [PubMed: 19755532]

Curley JP, Keverne EB. Genes, brains and mammalian social bonds. Trends in Ecology & Evolution. 2005; 20(10):561–7. [PubMed: 16701435]

Depue RA, Morrone-Strupinsky JV. A neurobehavioral model of affiliative bonding: implications for conceptualizing a human trait of affiliation. Behavioral and Brain Sciences. 2005; 28(3):313–95. [PubMed: 16209725]

Dezecache G, Dunbar RIM. Sharing the joke: the size of natural laughter groups. Evolution & Human Behaviour. 2012; 33(6):775–779.

Dunbar RIM. Functional Significance of Social Grooming in Primates. Folia Primatologica. 1991; 57:121–31.

Dunbar RIM. The social brain hypothesis. Evolutionary Anthropology: Issues, News, and Reviews. 1998; 6(5):178–90.

Dunbar RIM. The Social Brain: Mind, Language, and Society in Evolutionary Perspective. Annual Review of Anthropology. 2003; 32:163–81.

Dunbar RIM. Mind the gap: or why humans aren't just great apes. Proceedings of the British Academy. 2008; 15:403–23.

Dunbar RIM. The social role of touch in humans and primates: behavioural function and neurobiological mechanisms. Neuroscience and Biobehavioral Reviews. 2010; 34:260–8. [PubMed: 18662717]

Dunbar RIM, Baron R, Frangou A, Pearce E, Leeuwen EJC, Van Partridge G, Stow J. Social laughter is correlated with an elevated pain threshold. Proceedings of the Royal Society B: Biological Sciences. 2012a; 279(1731):1161–7. [PubMed: 21920973]

Dunbar RIM, Kaskatis K, Macdonald I, Barra V. Performance of music elevates pain threshold and positive affect: implications for the evolutionary function of music. Evolutionary Psychology. 2012b; 10(4):688–702. [PubMed: 23089077]

Fabre-Nys C, Meller RE, Keverne EB. Opiate antagonists stimulate affiliative behaviour in monkeys. Pharmacology, Biochemistry, and Behavior. 1982; 16(4):653–9.

Getz LM, Chamorro-Premuzic T, Roy MM, Devroop K. The relationship between affect, uses of music, and music preferences in a sample of South African adolescents. Psychology Of Music. 2012; 40(2):164–78.

Guemple, L., editor. Alliance in Eskimo Societies. New York, NY, US: American Ethnological Society; 1971.

Hamilton MJ, Milne BT, Walker RS, Burger O, Brown JH. The complex structure of hunter-gatherer social networks. Proceedings of the Royal Society B: Biological Sciences. 2007; 274(1622):2195–202. [PubMed: 17609186]

Hill R, Dunbar RIM. Social network size in humans. Human Nature. 2003; 14(1):53–72. [PubMed: 26189988]

Holt-Lunstad J, Smith TB, Layton JB. Social Relationships and Mortality Risk: A Meta-analytic Review. PLoS Medicine. 2010; 7:e1000316. [PubMed: 20668659]

Hove M, Risen J. It's all in the timing: Interpersonal synchrony increases affiliation. Social Cognition. 2009; 27(6):949–60.

Hsu DT, Sanford BJ, Meyers KK, Love TM, Hazlett KE, Wang H, Zubieta J-K. Response of the μ-opioid system to social rejection and acceptance. Molecular Psychiatry. 2013; 18(11):1211–7. [PubMed: 23958960]

Huron, D. Is music an evolutionary adaptation?. The cognitive neuroscience of music. Peretz, I.; Zatorre, R., editors. New York, NY, US: Oxford University Press; 2003. p. 57-75.

Kalender B, Trehub SE, Schellenberg E. Cross-cultural differences in meter perception. Psychological Research. 2013; 77(2):196–203. [PubMed: 22367155]

Kawakami A, Furukawa K, Katahira K, Okanoya K. Sad music induces pleasant emotion. Frontiers in Psychology. 2013; 4

Keverne EB, Martensz ND, Tuite B. Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships. Psychoneuroendocrinology. 1989; 14(1&2):155–61. [PubMed: 2525263]

Kirschner S, Tomasello M. Joint music making promotes prosocial behavior in 4-year-old children. Evolution and Human Behavior. 2010; 31(5):354–64.

Laiho S. The Psychological Functions of Music in Adolescence. Nordic Journal Of Music Therapy. 2004; 13(1):47–63.

Launay J, Dean RT, Bailes F. Synchronization can influence trust following virtual interaction. Experimental Psychology. 2013; 60(1):53–63. [PubMed: 22935329]

Launay J, Dean RT, Bailes F. Synchronising movements with the sounds of virtual partner enhances partner likeability. Cognitive Processing. 2014; 15(4):491–501. [PubMed: 24805849]

Launay J, Grube M, Stewart L. Dysrhythmia: A specific congenital rhythm perception deficit. Frontiers in Psychology: Auditory Cognitive Neuroscience. 2014; 5

Lehmann J, Korstjens AH, Dunbar RIM. Group size, grooming and social cohesion in primates. Animal Behaviour. 2007; 74(6):1617–29.

Lehmann, J.; Lee, PC.; Dunbar, RIM. Unravelling the evolutionary function of communities. Lucy to language: the benchmark papers. Dunbar, RIM.; Gamble, C.; Gowlett, JAJ., editors. Oxford, UK: Oxford University Press; 2014. p. 245-76.

Machin AJ, Dunbar RIM. The brain opioid theory of social attachment: a review of the evidence. Behaviour. 2011; 148(9):985–1025.

Machin AJ, Dunbar RIM. The brain opioid theory of social attachment: a review of the evidence. Behaviour. 2011; 148(9):985–1025.

Mackinnon A. A short form of the Positive and Negative Affect Schedule: Evaluation of factorial validity and invariance across demographic variables in a community sample. Personality and Individual Differences. 1998; 27(3):405–16.

Martel FL, Nevison CM, Simpson MJ, Keverne EB. Effects of opioid receptor blockade on the social behavior of rhesus monkeys living in large family groups. Developmental Psychobiology. 1995; 28(2):71–84. [PubMed: 8529786]

Meller RE, Keverne EB, Herbert J. Behavioural and endocrine effects of naltrexone in male talapoin monkeys. Pharmacology, Biochemistry, and Behavior. 1980; 13(5):663–72.

Miranda D, Gaudreau P. Music listening and emotional well-being in adolescence: A person- and variable-oriented study. Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology. 2011; 61(1):1–11.

Pearce E. Modelling mechanisms of social network maintenance in hunter–gatherers. Journal of Archaeological Science. 2014; 50:403–413. [PubMed: 25214706]

Pearce E, Launay J, Dunbar RIM. The Ice-breaker Effect: Singing together mediates fast social bonding. Royal Society Open Science. in press.

Pearce, E.; Shuttleworth, A.; Grove, M.; Layton, R. The costs of being a high latitude hominin. The Lucy Project: Benchmark Papers. Dunbar, RIM.; Gamble, C.; Gowlett, J., editors. Oxford, UK: Oxford University Press; 2014.

Reddish P, Fischer R, Bulbulia J. Let's dance together: synchrony, shared intentionality and cooperation. PloS One. 2013; 8(8):e71182. [PubMed: 23951106]

Sebanz N, Knoblich G. Prediction in joint action: What, when and where? Topics in Cognitive Science. 2009; 1(2):353–67. [PubMed: 25164938]

Shteynberg G, Hirsh JB, Galinsky AD, Knight AP. Shared attention increases mood infusion. Journal of Experimental Psychology: General. 2014; 143(1):123–30. [PubMed: 23317087]

Sloboda JA, Davidson JW, Howe MJ. Is everyone musical? -Target paper. The Psychologist England. 1994:349–354.

Sullivan PJ, Rickers K, Gammage K. The effect of different phases of synchrony on pain threshold. Group Dynamics: Theory, Research and Practice. 2014; 18(2):122–8.

Sullivan P, Rickers K. The effect of behavioral synchrony in groups of teammates and strangers. International Journal of Sport and Exercise Psychology. 2013; 11(3):286–91.

Tarr B, Launay J, Dunbar RIM. Music and social bonding: "self-other"• merging and neurohormonal mechanisms. Frontiers in Psychology. 2014; 5:1–10. [PubMed: 24474945]

Umberson D, Montez JK. Social Relationships and Health: A Flashpoint for Health Policy. Journal of Health and Social Behavior. 2010; 51(1 suppl):S54–S66. [PubMed: 20943583]

Umberson D, Montez JK. Social relationships and health a flashpoint for health policy. Journal of Health and Social Behavior. 2010; 51(1 suppl):S54–S66. [PubMed: 20943583]

Valdesolo P, Desteno D. Synchrony and the social tuning of compassion. Emotion. 2011; 11(2):262–6. [PubMed: 21500895]

Vickhoff B, Malmgren H, Aström R, Nyberg G, Ekström S-R, Engwall M, Jörnsten R. Music structure determines heart rate variability of singers. Frontiers in Psychology. 2013; 4:334. [PubMed: 23847555]

Whallon R. Social networks and information: Non-"utilitarian" mobility among hunter-gatherers. Journal of Anthropological Archaeology. 2006; 25:259–70.

Wiltermuth SS, Heath C. Synchrony and cooperation. Psychological Science. 2009; 20(1):1–5. [PubMed: 19152536]

Wolf W, Launay J, Dunbar RIM. Joint attention, shared goals, and social bonding. British Journal of Psychology. 2015; doi: 10.1111/bjop.12144

Zhou W-X, Sornette D, Hill RA, Dunbar RIM. Discrete hierarchical organization of social group sizes. Proceedings of the Royal Society B: Biological Sciences. 2005; 272(1561):439–44. [PubMed: 15734699]

Zubieta JK, Smith YR, Bueller JA, Xu Y, Kilbourn MR, Jewett DM, Stohler CS. Regional mu opioid receptor regulation of sensory and affective dimensions of pain. Science. 2001; 293(5528):311–5. [PubMed: 11452128]

Zubieta J-K, Ketter TA, Bueller JA, Xu Y, Kilbourn MR, Young EA, Koeppe RA. Regulation of human affective responses by anterior cingulate and limbic mu-opioid neurotransmission. Archives of General Psychiatry. 2003; 60(11):1145–53. [PubMed: 14609890]
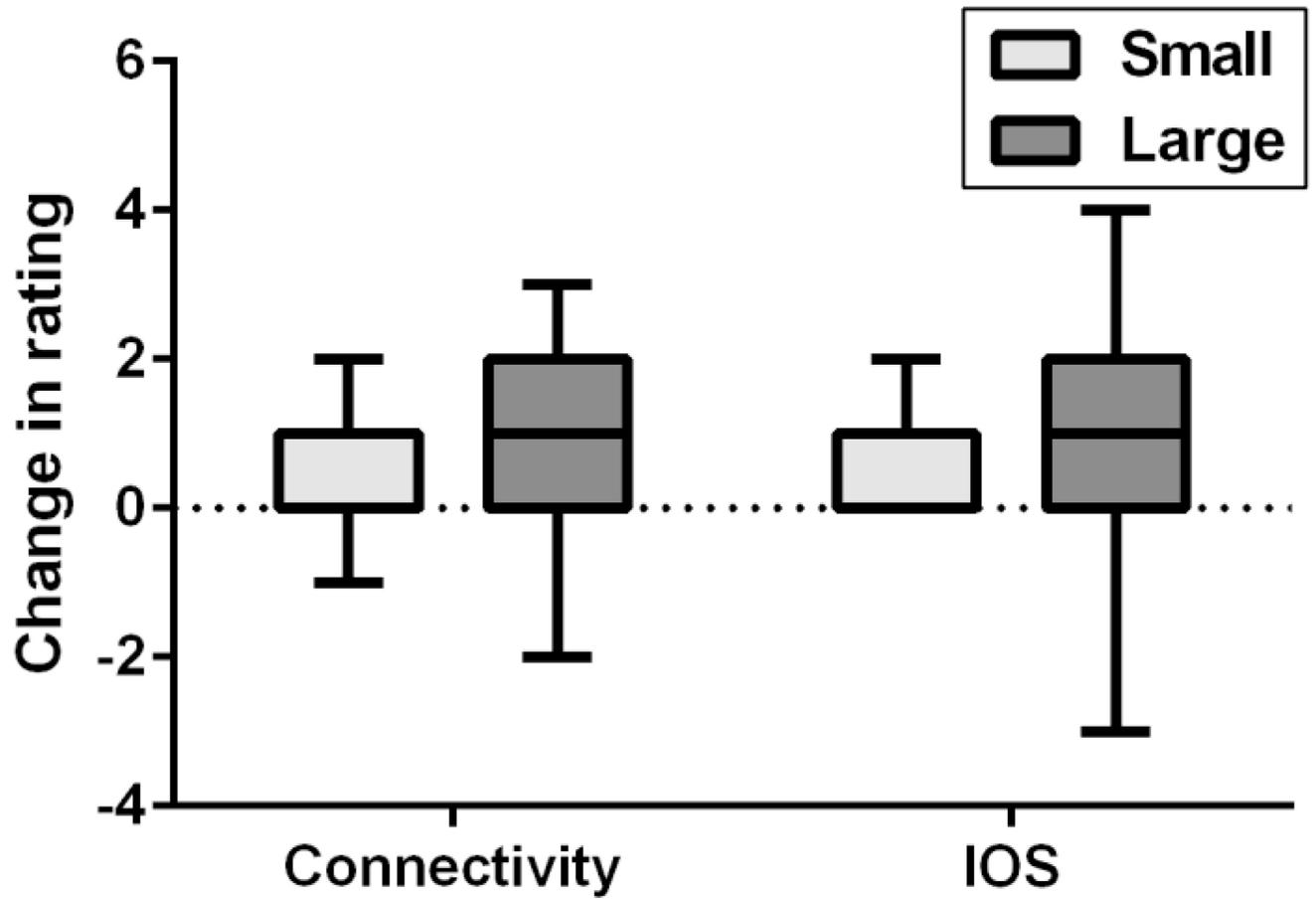
Figure 1.
Tukey boxplot showing change in IOS and connectivity ratings following singing in small and large choir conditions.
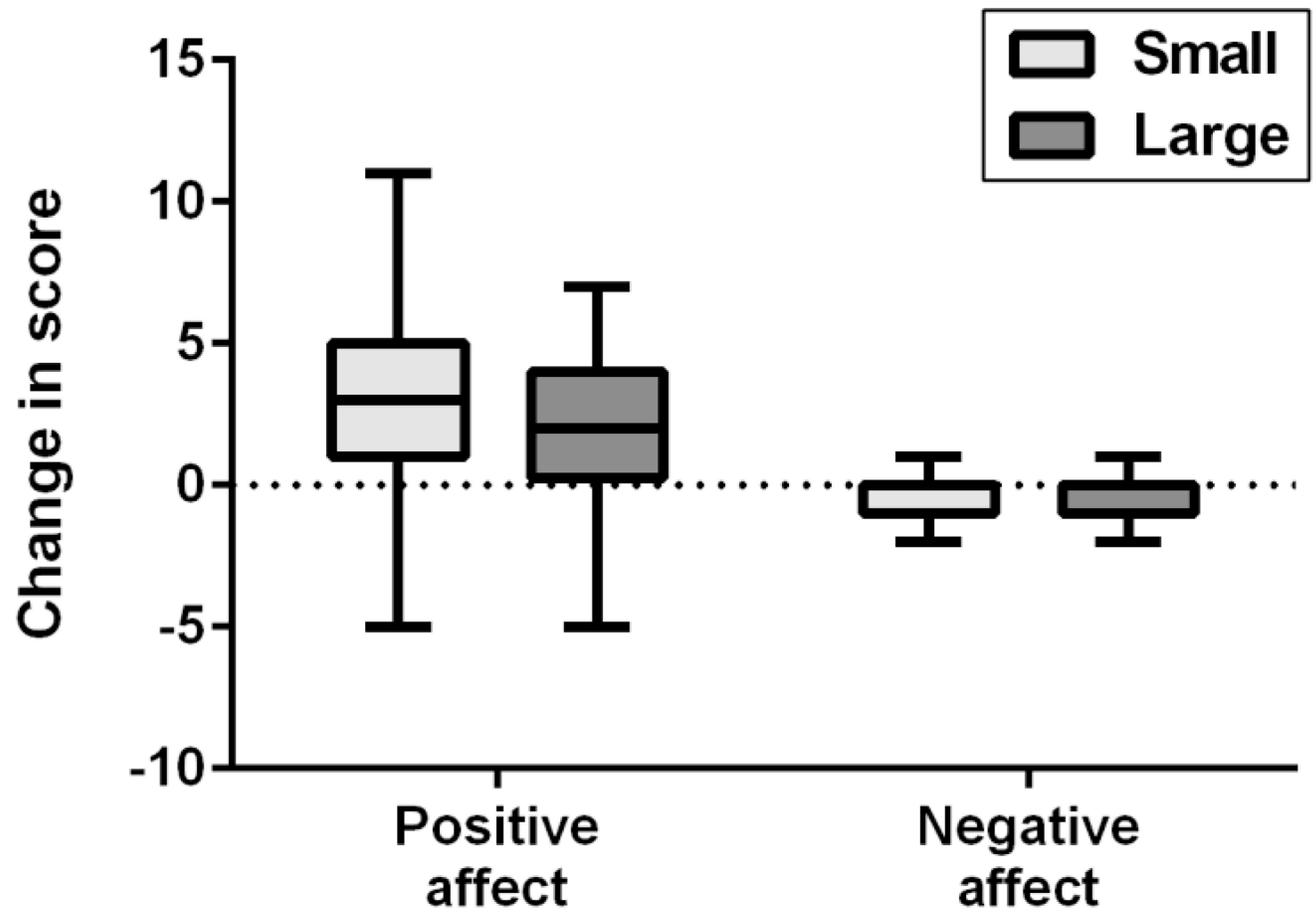
Figure 2.
Tukey boxplot showing change in positive and negative affect, as measured by PANAS scale, following singing in small and large choir conditions.
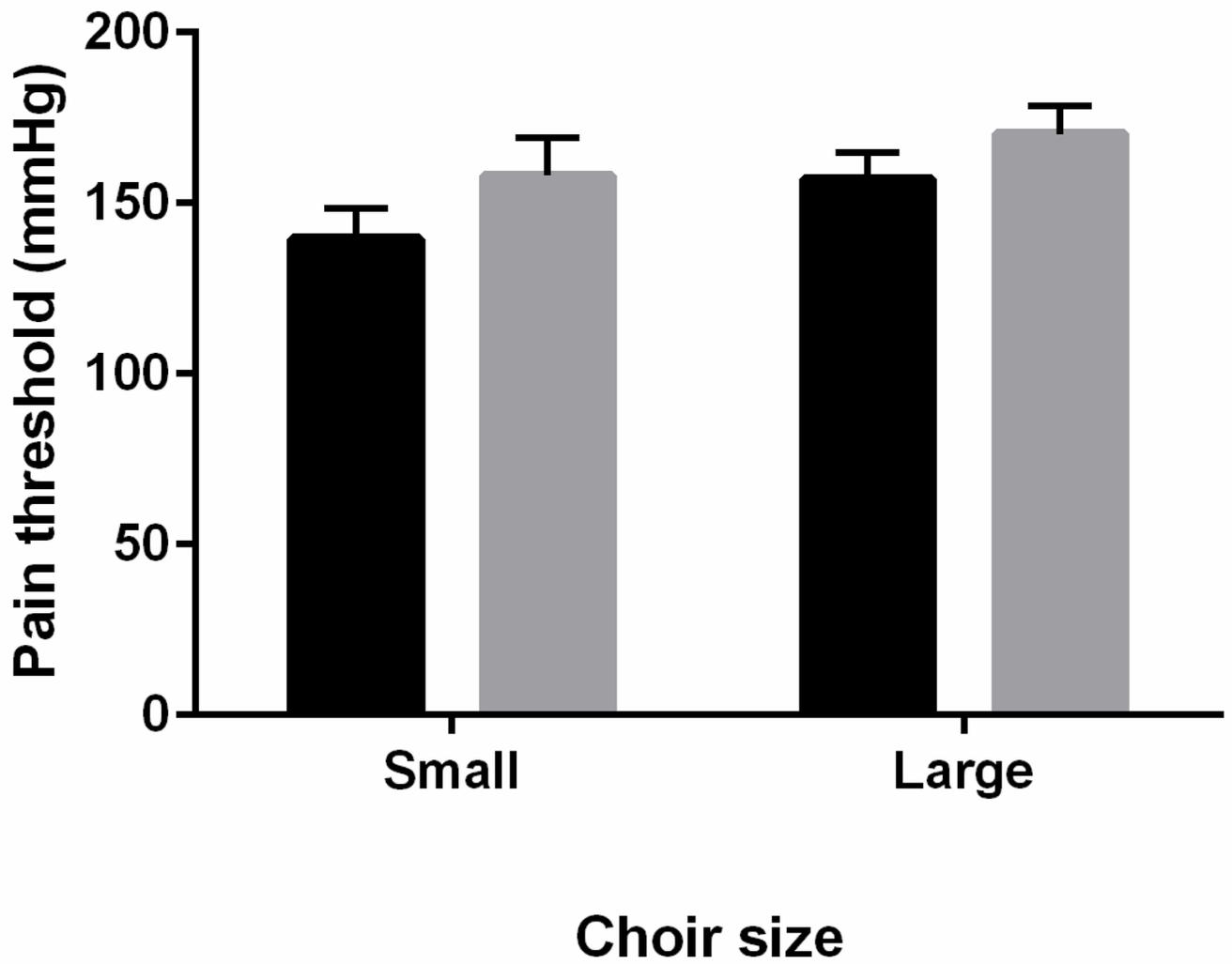
Figure 3.
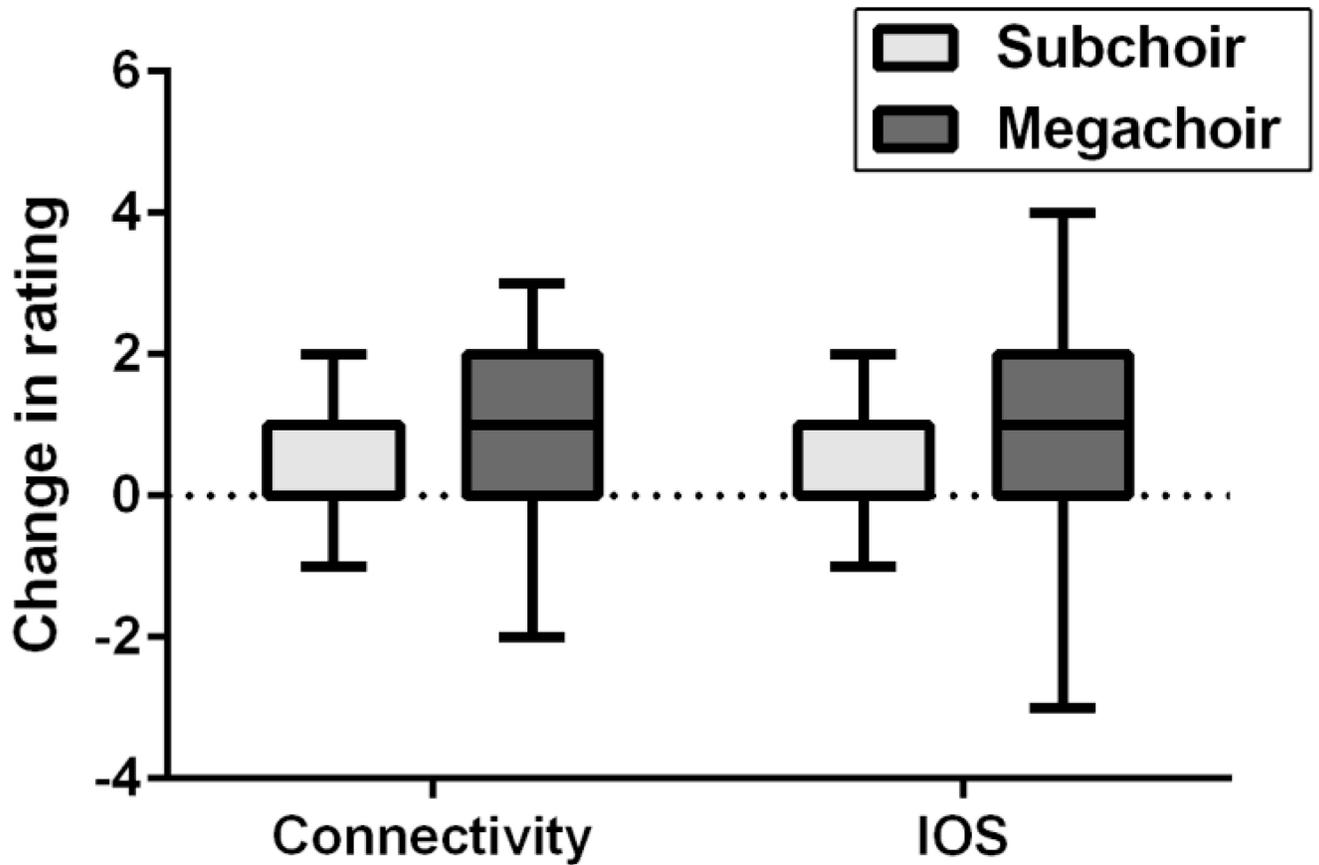Pain thresholds before (black bar) and after (grey bar) singing in small and large choir conditions.

Figure 4.
Tukey boxplot showing change in IOS and connectivity ratings made with reference to the subchoir and megachoir.

Table 1

Descriptive statistics (Mean and SD) for all self-report measures.

	Small choir (*n* = 133)		Subchoir (*n* = 80)		Megachoir (*n* = 80)	
	Before *M* (*SD*)	**After *M* (*SD*)**	**Before *M* (*SD*)**	**After *M* (*SD*)**	**Before *M* (*SD*)**	**After *M* (*SD*)**
IOS	3.67 (1.53)	4.50 (1.52)	4.00 (1.56)	4.36 (1.55)	2.71 (1.37)	3.95 (1.57)
Connectivity	4.29 (1.38)	5.08 (1.20)	4.39 (1.25)	4.80 (1.29)	3.49 (1.25)	4.61 (1.15)
Positive Affect	14.4 (4.16)	17.6 (4.43)	-	-	15.9 (4.12)	18.0 (4.11)
Negative Affect	6.11 (2.00)	0.50 (1.21)	-	-	5.61 (1.23)	5.24 (0.97)