View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by White Rose Research Online

This is a repository copy of *Promoting prosociality in Colombia: Is music more effective than other cultural interventions?*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/176048/

Version: Published Version

Article:
Cespedes-Guevara, J. and Dibben, N. orcid.org/0000-0002-9250-5035 (2021) Promoting prosociality in Colombia: Is music more effective than other cultural interventions? Musicae Scientiae. ISSN 1029-8649

https://doi.org/10.1177/10298649211013505

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

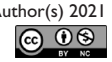
White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Check for updates

Article

MUSICÆ SCIENTIÆ

Musicae Scientiae
1–26
© The Author(s) 2021

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10298649211013505
journals.sagepub.com/home/msx


Promoting prosociality in Colombia: Is music more effective than other cultural interventions?

Julian Cespedes-Guevara
Universidad Icesi, Colombia

Nicola Dibben
University of Sheffield, UK

Abstract

This article reports a two-part study into the prosocial impacts of third sector cultural activities with children and adolescents in impoverished and violence-stricken urban neighbourhoods in Cali, Colombia. First, a year-long field study set out to compare a pre-existing music-training programme with a dance-training programme and a football-training programme with 9–14 year olds, to determine the extent to which each affords the development of empathic attitudes and prosocial behaviours. The music and dance programmes produced few significant changes in participants' empathy or prosociality, and there were few significant differences between the empathy and prosociality of the participants in the two groups. Participant dropout prevented comparison with the football-training programme. Second, an interview study was used to understand the place of prosociality in the aims and work of policymakers, funders and third-sector practitioners running cultural activities for social impacts in the Cali region. The study revealed that the organisations aimed to achieve individual and social transformation by creating the conditions for transformation, evidenced as positive outcomes. Neither the measures used by the organisations themselves nor the psychosocial constructs of prosociality and empathy used by the researchers adequately evidenced some of the intended outcomes, such as enabling individuals to build a *life project*, practising and sustaining social inclusion and transforming communities, nor a path from individual to social transformation. Differences between the structure of cultural activities and their associated values meant that different activities were believed to lend themselves to social transformation more or less well. This highlights the need for critically reflective, co-constructed research using a fuller range of constructs that can capture the outcomes of these programmes for both individuals and groups.

Keywords

Music making, prosocial, empathy, non-formal learning, music, dance, football

Corresponding author:
Julian Cespedes-Guevara, Universidad Icesi, Colombia.
Email: jcespedes@icesi.edu.co

Music and other cultural activities (artistic, sporting and intellectual) are commonly used to achieve social impacts – changes for individuals, groups, societies and social structures.[1] However, there is little understanding of the relative effectiveness of different cultural interventions nor how meta-level constructs can be operationalised in order to make them amenable to systematic study: impacts happen at different relational proximities (the individual, immediate family and friends, local community and society at large), and timescales (from during the activity, to those over a lifetime). Understanding of the social impacts of cultural activities, including music, is currently fragmented and lacks an overarching taxonomy or framework that can encompass these characteristics (Sloboda, 2019). The importance of understanding how different cultural activities bring about which social impacts is evident when faced with on-the-ground decisions about cultural provision premised on the ability of culture to achieve social impacts.

Such is the place of cultural activities[2] for social impacts in Colombia, which we take as our case study. As a consequence of one of the longest-running civil wars in modern times, Colombia has more than eight million victims of conflict and the largest internally displaced population in the world, resulting in tensions due to financial, social, economic and cultural inequality and exclusion. The Colombian Culture Ministry, non-governmental organisations (NGOs) and independent organisations use cultural activities to achieve social and economic development – see Colombia's National Development Plan, 2014–2018 (DNP Colombia, 2015). Our research investigated the social impacts of artistic- and sports-training programmes with young people in Cali, Colombia, and the differing priorities and interests of academics and practitioners connected to these programmes.

Current models of the social impacts of the arts emphasise the role of cultural activities in nurturing prosociality (helping, caring and empathising with others) (Broadwood et al., 2012; Tay et al., 2018; Van de Vyver et al., 2019). In these models, prosociality is associated with both individual and social well-being (overcoming inequality and humanitarian crises), although there is an implicit leap from individual prosociality to larger social outcomes.

Prosocial behaviour encompasses behaviours such as helping, comforting, sharing, cooperating, reassuring and defending, where the intention is to help others. Prosocial behaviour can be measured by self-report questionnaires (e.g., Auné et al., 2016) or through situational behavioural tasks (e.g., Kirschner & Tomasello, 2010). Empathy is an important construct linked to prosociality, and refers to understanding and sharing the feelings of another person, both as a trait (dispositional empathy) and as a response to a situation (situational empathy) (Eisenberg et al., 2010).

1 The role of artistic and sporting activities in creating more cohesive and inclusive societies is formalised in a UN Resolution (A/C.2/70/L.59 at the 70th session of the UN General Assembly, 14 December 2015), which states that the economic and social dimensions of poverty can be addressed through cultural heritage and the cultural and creative industries. Specifically, Point 1 of the Resolution emphasises "the role of culture as an enabler of sustainable development that provides people and communities with a strong sense of identity and social cohesion", and Point 9j recognises the need "To mobilize culture as a vehicle to foster tolerance, mutual understanding, peace and reconciliation in the context of conflict-prevention, conflict-resolution and peacebuilding processes".

2 We use the word "culture" throughout this article to denote artistic, sporting and intellectual activities, rather than culture in the sense of individual enrichment, and/or ways of living. These three meanings are intertwined such that they communicate an ideal in which people encounter a good way to live, which is expressed in their ways of living and through which individuals can flourish.

Some of the evidence for enhanced empathy and prosocial behaviour with arts engagement comes from laboratory and quasi-experimental studies. Year-long programmes of interactive musical games (Rabinowitch et al., 2013) and acting (Goldstein & Winner, 2012) have been found to enhance empathy relative to control conditions. Moreover, association has been reported between starting instrumental lessons at an early age and sociability (Kawase et al., 2018), and between enhanced prosociality and a 10-month course of standard music training, albeit for those who started the programme with lower scores at the outset and using a self-report measure rather than a behavioural measure (Schellenberg et al., 2015). In addition, an experiment comparing joint music making with an equivalent non-musical play session showed increased helping and cooperation behaviours among four-year-olds immediately after the session (Kirschner & Tomasello, 2010).

An important question is whether and how cultural activities can develop out-group prosociality, because prosociality is more likely to occur in relation to members of the same group (Raabe & Beelmann, 2011). One study to have investigated prosocial intentions and empathy of children engaging in kindness-focused participatory arts programmes found an association between engagement in the programme and prosocial intentions, but not empathy, and increased prosocial intentions towards (fictional) out-group members in competitive contexts that lasted six months (Van de Vyver et al., 2019).

Little if any research has investigated whether there is a difference in the type, longevity or extent of impact on prosociality according to the type of cultural activity involved. The few comparative studies that have been reported investigate other outcomes and indicate broad similarities as regards physical, psychological and social benefits for amateur knitters and musicians (Lamont & Ranaweera, 2019), and improvements in adult mental health from creative writing or singing in choirs (Williams et al., 2019). Nonetheless, there is reason to think that the weighting of different outcomes will differ according to the characteristics of the activity. Different activities afford differences in the rapidity with which adult groups form bonds, and the number of bonds they form, in singing compared to crafts and creative writing classes (Pearce et al., 2015) and dance compared with conversation (Robertson et al., 2017). Thus, structural features of the activity afford different kinds of engagement and potentially different outcomes but there have been no direct comparisons of pre-existing cultural activities to test this. Interestingly, a large UK survey found prosocial behaviour to be more positively associated with arts engagement than other variables including engagement with sport, but this included both participatory and presentational forms of arts engagement, which makes comparison with other studies difficult (Van de Vyver & Abrams, 2018).

As indicated by the foregoing, insight into the influence of participatory cultural engagement on prosocial behaviour and human flourishing is limited in specific ways. First, comparative empirical data is scarce. Previous studies tend to focus on single activities, and there is a lack of focus on explicit and consistent outcome measures that might allow both generalisability and comparison within and between cultural activities. A comparative approach, although not free of biases, reduces the likelihood of influence from a vested interest in, and assumptions about particular cultural domains (Baker, 2014; Sloboda, 2019). Moreover, although there is evidence of specific psychosocial and emotional processes underlying prosocial behaviour, many studies are laboratory or quasi-experimental studies and there are minimal data from field studies of pre-existing intervention programmes (evaluations of existing programmes have been criticised for their alignment with advocacy rather than research: Baker et al., 2018). Second, there is limited understanding of the aims and values of those providing the interventions and therefore a potential mismatch between what researchers measure and what cultural programmes set out to change.

Given this deficit, we designed a two-part study that aimed (i) to test whether prosociality and empathy increase with participation in pre-existing arts and sports interventions, and if the increase differs across activity type; (ii) regardless of the outcome of (i), to understand which outcomes are intended by pre-existing arts and sports interventions and the relevance of prosociality and empathy to these. Our project engaged with Colombian providers of participatory cultural activities in the form of music training, football training and dance training, which had as their primary goal social impacts intended to promote both individual and societal human flourishing. Ethical approval for the studies was granted by the research ethics committees of the University of Sheffield and Icesi University.

Study 1: Comparative field study

Our primary research question was to what extent cultural activities differ in their ability to increase prosociality and empathy. To this end we compared pre-existing programmes aimed at achieving social outcomes in music, dance and sport, since these differ, for example, in the opportunities they offer for physical exertion and competition. We investigated their work with children aged 9–14, because this group comprises the target population for these NGOs and is underrepresented in existing research. We used established measures of the psychosocial constructs of prosociality and empathy. As in the other studies reporting psychosocial interventions within which we contextualise our research, the data were gathered outside the moment of the activity itself.

Method

Design. This comparative study used a non-experimental longitudinal design, with type of training (music, dance or football) as the independent variable, and empathy and prosocial behaviour as the dependent variables. Participants were contacted and assessed fewer than three months after they began training in the respective cultural activities, and one year later.

Participants. Participants were recruited from NGOs that provide social intervention programmes to children and young people from impoverished and violence-stricken urban neighbourhoods in Cali, Colombia. Musicians were recruited from two similar NGOs that provide classical music instrumental training: an industry-sponsored youth symphony orchestra and an arts organisation ($n = 20$). Dancers were recruited from the arts organisation ($n = 20$), and footballers were recruited from three NGOs: an industry-sponsored football club (same organisation as the Youth Symphony Orchestra) and two independent football clubs ($n = 24$). These NGOs were chosen because they describe themselves as educational institutions that provide non-formal musical education, or as cultural sector organisations whose purpose is to provide social well-being to vulnerable and disadvantaged children and young people (aged 2–20), living in areas of low economic income and high criminality.

Unfortunately, although most of the children we originally contacted continued in their programmes after one year ($N = 60$), we faced a number of difficulties that meant that we were only able to contact and test 11 children from the original music-training group, nine from the dance group, and one from the football group. The data from the football group were therefore excluded from the analysis reported in the following.

Measures. The Spanish version of the Toronto Empathy Questionnaire (TEQ) was used to measure Trait Empathy, defined as the ability to gain "accurate affective insight into the feeling state

of another" (Spreng et al., 2009, p. 68). It consists of 16 items with statements such as "When someone else is feeling excited, I tend to get excited too", rated on a Likert-type scale from 0 (*never*) to 4 (*always*). Responses are summed to generate a total score out of 64; higher scores indicate more empathy.

Situational empathy was measured using the empathy for pain task (Decety et al., 2012), in which the participant is shown short videos portraying three types of situations: intentional harm scenes, where a person causes pain to another deliberately, accidental harm scenes and control scenes where no harm is done. After watching each video, the participant is asked whether they think the act was intentional (cognitive empathy), how sad they are for the victim (empathic concern), how upset they feel (personal distress), whether they blame the perpetrator (moral blame) and whether they think the perpetrator deserves punishment (moral evaluation). Cognitive empathy is scored on a scale of 0 (*no intention to cause harm*) to 100 (*intention to cause harm*). Empathic concern, personal distress, moral blame and moral evaluation are rated on a scale from −12 to 12, where higher scores indicate more affective empathy and harsher moral attitudes.

Prosocial behaviour was measured using a situational task based on a procedure originally used by Kokal et al. (2011), where one of the researchers "accidentally" drops 12 objects (pencils or sheets of paper), and the number of objects collected by the participant is counted. This technique allowed us to observe helpful behaviours directly, rather than reported by the participants in a questionnaire.

Procedure. Agreement and informed consent were obtained from the NGO directors and the children's guardians. These NGOs recruited children over the course of several weeks, and there were delays in the returning of signed consent forms, so the first assessment included children who were up to three months into their respective programmes.

Participants were tested individually by pairs of researchers, during their regular classes; the child was accompanied to a room and asked to sit at a desk with a laptop that showed the empathy for pain task. Participants were asked to read the instructions and answer the questions on screen. Then, participants were given a sheet of paper with the TEQ, and the prosocial behaviour task was carried out (one researcher dropped pencils on the floor, while the other wrote down the number of pencils picked up by the participant). After this, participants completed the questionnaire, and were thanked for their participation. The procedure lasted around 20 minutes in total. During the second phase of assessments, the procedure was identical except that the researcher dropped 12 sheets of paper instead of pencils.

Results

At the first assessment point, the sample consisted of 40 children from 9–14 years of age ($M = 11.3$, $SD = 1.51$) who had recently started their training (< 3 months) in either music ($n = 20$; $M_{age} = 11.5$, $SE = 1.47$) or dance ($n = 20$, $M_{age} = 11.1$, $SE = 1.55$). At the second assessment point, we collected data from 20 children from the original sample ($M_{age} = 12.70$, $SE = 1.26$): 11 children from the music programme ($M_{age} = 13.00$, $SE = 1.18$) and 9 children from the dance programme ($M_{age} = 12.33$, $SE = 1.32$). There were no significant differences between the mean ages of the two groups at the first assessment, $t(38) = 8.37$, $p = .408$, or at the second, $t(18) = 1.19$, $p = .20$.

An initial analysis using the Shapiro–Wilk statistic revealed that data from the three dependent measures were not normally distributed, so we analysed the data using non-parametric tests.

There were no significant differences between the Trait Empathy scores of the participants in the music and dance groups in either the initial or the second assessment, as indicated by Mann–Whitney tests of the TEQ scores (both $p > .05$; Tables 1 and 2). And although both groups had slightly lower scores in the second assessment, a Wilcoxon signed-rank test indicated that these scores were not significantly different from the groups' own scores at the first point of measurement (both $p > .05$; Table 3 and Figure 1).

In the prosocial behaviour task, the dance group had higher scores than the music group in both assessments, but the Mann–Whitney tests indicated that these differences were nonsignificant (both $p > .05$; Tables 1 and 2). In the second assessment, the music group had lower scores than in the first, and the dance group had higher scores, but the Wilcoxon signed-rank tests indicated that none of these changes was statistically significant for any of the groups (both $p > .05$; Table 3 and Figure 2).

The mean cognitive empathy scores in the situational empathy measure (empathy for pain task) ranged from 64.25% to 88.37%, indicating that on most stimuli, the participants identified correctly whether the portrayed harm was intentional or not. The reliability of this measurement technique is also confirmed by the fact that in both assessments, the statistical tests showed a pattern of results consistent with the intentionality level of the situations portrayed in the stimuli. Thus, the intentional harm situations were associated with the highest scores for cognitive empathy, empathic concern, personal distress and with harsher moral evaluations, followed by the accidental harm situations and, finally, by the neutral situations (Table 4).

In the first assessment, in those stimuli where the empathy for pain task portrayed intentional harm situations, the Mann–Whitney tests showed that the dance group had significantly higher scores than the music group for cognitive empathy ($p = .006$), empathic concern ($p = .017$) and moral evaluation (indicating harsher judgements) ($p = .024$, see Table 1). However, in the second assessment, the music group caught up with the dance group, so to speak, as indicated by Mann–Whitney tests showing that there were no significant differences between the groups for any of these measures one year later (Table 2). When we compared the scores of each group in the first and the second assessments, we found only one significant change: a Wilcoxon signed-rank test indicated that for the music group, the scores for empathic concern were significantly higher in the second assessment (mean rank = 6.40)[3] than the first assessment (mean rank = 2.00), $Z = 2.76$, $p = .006$. All other comparisons were nonsignificant (Table 3).

When the empathy for pain task portrayed accidental harm situations, the Mann–Whitney tests did not show any significant differences between the scores of the music and the dance groups in the first or second assessment (all $p > .05$, Tables 1 and 2). A comparison of the first and second assessments using Wilcoxon signed-rank tests showed that the only significant change was found in the dance group, which had significantly lower scores for empathic concern in the second assessment (mean rank = 2.00), than in the first assessment (mean rank = 5.38), $Z = -2.43$, $p = .015$ (Table 3).

We did not find any differences between the scores of the music and dance groups in any of the assessments when the empathy for pain task portrayed neutral situations (all $p > .05$; see Tables 1 and 2 for details). When comparing the scores from the first and second assessments,

3 In nonparametric statistical tests, raw scores are ordered from lowest to highest, and the lowest score is assigned a rank of 1, the next highest score is assigned a rank of 2, and so on. The mean rank is the average of assigned ranks and is considered the rank with the most common frequency (Field, 2013). Hence, the mean rank is roughly equivalent to the mean in parametric tests.

Table 1. Descriptive statistics and results of Mann–Whitney tests of dependent measures in the first assessment.

Measure		Programme N_{Music} = 20 N_{Dance} = 20	Mean (DE)	Confidence Intervals 95%	Mean Rank	Z	*p*
Toronto Empathy Questionnaire		Music	**36.70** (7.46)	33.21–40.19	17.72	1.50	.133
		Dance	**40.50** (9.24)	36.18–44.82	23.38		
Prosocial behaviour Task		Music	**4.95** (3.91)	3.12–6.78	18.68	1.00	.318
		Dance	**6.35** (4.31)	4.33–8.37	22.32		
Cognitive Empathy	Intentional Harm	Music	**67.72** (23.41)	56.77–78.68	15.50	–2.77	.006*
		Dance	**73.33** (18.96)	64.46–82.21	25.58		
	Accidental Harm	Music	**77.00** (35.68)	60.31–93.70	20.70	–.11	.914
		Dance	**77.44** (87.02)	61.13–93.75	20.30		
	Neutral Situation	Music	**69.50** (37.54)	51.93–87.07	23.02	–1.37	.172
		Dance	**51.10** (40.19)	32.29–69.90	17.98		
Empathic Concern	Intentional Harm	Music	**4.40** (2.51)	3.23–5.57	16.20	–2.39	.017*
		Dance	**6.32** (3.15)	4.84–7.79	24.80		
	Accidental Harm	Music	**1.50** (4.04)	−0.39 to 3.39	17.55	1.60	.110
		Dance	**3.47** (3.66)	1.76–5.18	23.45		
	Neutral situation	Music	**–4.37** (3.57)	−6.04 to −2.69	20.30	.11	.913
		Dance	**−4.36** (3.43)	−5.97 to −2.76	20.70		
Personal Distress	Intentional Harm	Music	**5.01** (5.67)	3.59–6.43	17.05	1.92	.055
		Dance	**6.18** (3.57)	4.51–7.86	23.95		
	Accidental Harm	Music	**0.58** (3.83)	−1.21 to 2.37	18.38	1.15	.250
		Dance	**2.14** (4.12)	0.21–4.074	22.62		
	Neutral situation	Music	**−4.12** (4.13)	−6.05 to −2.18	20.98	–.26	.795
		Dance	**−4.47** (3.58)	−6.14 to −2.79	20.02		
Moral Evaluation	Intentional Harm	Music	**4.16** (3.49)	2.53–5.79	16.38	–2.26	.024*
		Dance	**6.19** (3.53)	4.54–7.85	24.63		
	Accidental Harm	Music	**−1.34** (4.47)	−3.43 to 0.75	18.95	.84	.401
		Dance	**−0.25** (4.67)	−2.44 to 1.93	22.05		
	Neutral situation	Music	**−4.37** (3.98)	−6.24 to −2.51	21.90	–.77	.444
		Dance	**−5.21** (3.49)	6.84 to −3.58	19.10		

Table 2. Descriptive statistics and results of Mann–Whitney Tests of dependent measures in the second assessment.

Measure		Programme N_{Music}= 11 N_{Dance}= 9	Mean (DE)	Confidence Intervals 95%	Mean Rank	Z	*p*
Toronto Empathy Questionnaire		Music	**33.45** (9.53)	27.05–39.86	10.18	.266	.824
		Dance	**35.44** (11.44)	26.65–44.23	10.89		
Prosocial behaviour task		Music	**3.27** (4.29)	0.39–6.16	8.41	1.81	.080
		Dance	**7.56** (5.48)	3.34–11.77	13.06		
Cognitive Empathy	Intentional Harm	Music	**69.14** (16.77)	57.87–80.41	8.27	1.88	.067
		Dance	**76.09** (8.23)	69.76–82.42	13.22		
	Accidental Harm	Music	**96.40** (13.95)	87.03–105.77	20.70	–.11	.914
		Dance	**79.39** (16.20)	66.87–91.78	20.30		
	Neutral situation	Music	**63.11** (36.00)	38.92–87.29	23.02	–1.37	.172
		Dance	**63.44** (29.27)	40.94–85.93	17.98		
Empathic Concern	Intentional Harm	Music	**6.15** (1.53)	5.12–7.18	9.68	.69	.503
		Dance	**6.55** (1.97)	5.04–8.07	11.50		
	Accidental Harm	Music	**1.05** (4.09)	−1.70 to 3.80	17.55	1.56	.110
		Dance	**−0.08** (3.90)	−3.08 to 2.91	23.45		
	Neutral situation	Music	**−6.96** (1.70)	−8.10 to −5.82	20.30	.11	.913
		Dance	**−6.71** (2.84)	−8.89 to −4.53	20.70		
Personal Distress	Intentional Harm	Music	**5.84** (2.16)	4.38–7.29	8.36	1.80	.080
		Dance	**7.42** (1.19)	6.50–8.33	13.11		
	Accidental Harm	Music	**0.24** (3.68)	−2.23 to 2.71	18.38	1.15	.250
		Dance	**0.31** (3.57)	−2.43 to 3.06	22.62		
	Neutral situation	Music	**−7.03** (1.88)	−8.30 to −5.78	20.98	–.26	.795
		Dance	**−5.92** (3.48)	−8.60 to −3.25	20.02		
Moral Evaluation	Intentional Harm	Music	**4.67** (4.61)	1.57–7.76	9.73	.65	.552
		Dance	**6.33** (2.17)	4.67–8.00	11.44		
	Accidental Harm	Music	**−3.71** (3.07)	−5.77 to −1.65	18.95	.84	.401
		Dance	**−2.15** (2.23)	−3.86 to −0.43	22.05		
	Neutral situation	Music	**−6.99** (1.76)	−8.17 to −5.81	21.90	–.77	.444
		Dance	**−6.65** (3.02)	−8.97 to −4.33	19.10		

Table 3. Wilcoxon tests of dependent measures: Assessment 1 versus Assessment 2.

Measure		Programme $N_{Music} = 11$ $N_{Dance} = 9$	Mean Rank		*Z*	*p*
			Assessment 1	Assessment 2		
Toronto Empathy Questionnaire		Music	4.50	5.63	.00	1.00
		Dance	3.88	5.90	–.83	.405
Prosocial behaviour task		Music	4.30	7.42	–1.03	.305
		Dance	4.14	7.00	–1.54	.123
Cognitive Empathy	Intentional Harm	Music	6.20	4.80	–.36	.721
		Dance	4.00	3.25	–.52	.600
	Accidental Harm	Music	4.30	6.70	–.61	.541
		Dance	4.00	4.67	–1.4	.161
	Neutral situation	Music	6.33	5.14	–.87	.386
		Dance	5.00	5.00	–1.48	.139
Empathic Concern	Intentional Harm	Music	2.00	6.40	–2.76	.006*
		Dance	4.50	3.80	–.85	.398
	Accidental Harm	Music	3.80	6.50	–.42	.678
		Dance	2.00	5.38	–2.43	.015*
	Neutral situation	Music	2.50	5.17	–1.82	.069
		Dance	6.00	4.29	–1.68	.092
Personal Distress	Intentional Harm	Music	8.40	4.00	–.80	.424
		Dance	5.00	2.75	–.11	.917
	Accidental Harm	Music	4.67	6.75	–.05	.959
		Dance	4.50	5.40	–.53	.594
	Neutral situation	Music	1.50	5.50	–2.10	.036*
		Dance	4.00	4.00	–.34	.735
Moral Evaluation	Intentional Harm	Music	6.57	5.00	–1.16	.248
		Dance	4.00	4.00	–.34	.735
	Accidental Harm	Music	4.50	5.14	–1.60	.110
		Dance	4.50	6.00	.53	.594
	Neutral situation	Music	3.83	5.58	–1.30	.192
		Dance	4.50	3.00	–.32	.752

we found that the music group had significantly lower scores for personal distress in the second assessment (mean rank = 1.50) than in the first (mean rank = 5.50), $Z = -2.10$, $p = .036$. All the other comparisons were nonsignificant (Table 3).

Discussion

A comparison of music- and dance-training groups in tests of prosociality and empathy showed few significant changes over a one-year period, and few significant differences between groups; taken together, these results suggest that neither the dance group nor the music group experienced changes in their empathic attitudes after one year of training. This conclusion is based on the observation that there was no consistent trend in the three measures we implemented; we did not find a uniform pattern of changes from the first to the second assessments in the trait empathy questionnaire, the situational prosocial behaviour task or the situational empathy task (see Figures 1–14). We did find some significant changes in scores for the

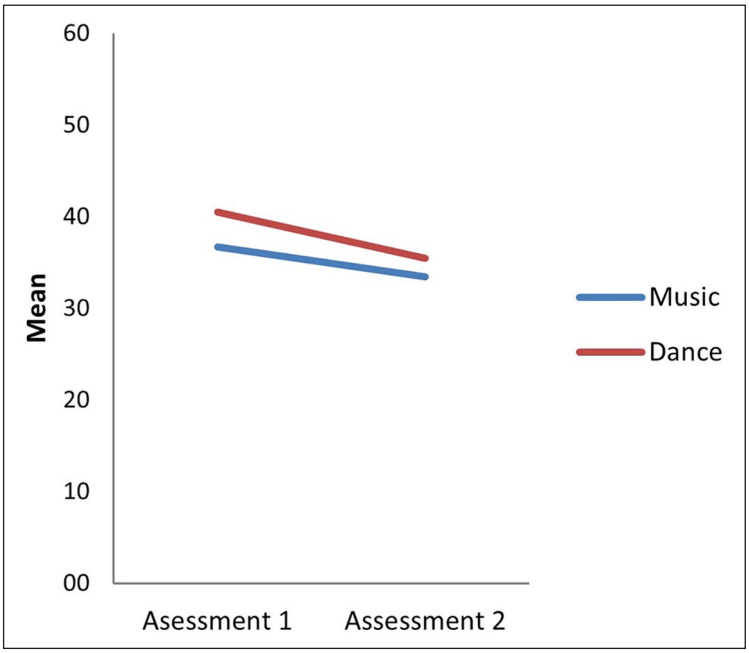


Figure 1. Mean scores: Toronto Empathy Questionnaire

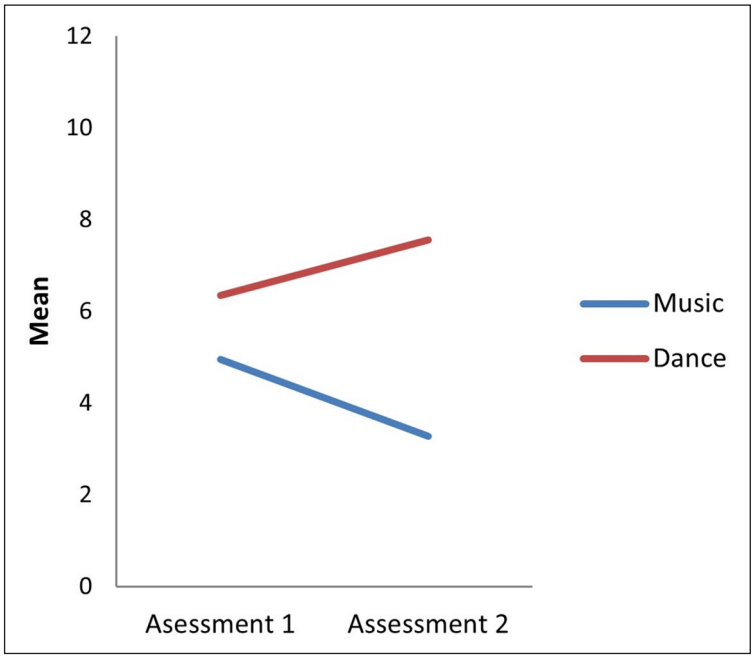


Figure 2. Mean scores: Prosocial Behaviour Task

Table 4. Participants' scores on the Empathy for Pain Task according to type of stimuli.

		Type of situation portrayed by the stimuli						Friedman Repeated Samples ANOVA		Wilcoxon Signed-rank Follow-up tests*		
		Intentional Harm		*Accidental Harm*		*Neutral*						
		Mean *(SD)*	Mean Rank	Mean *(SD)*	Mean Rank	Mean *(SD)*	Mean Rank	*Z*	*p*		*Z*	*p*
First Assessment $n = 40$	Cognitive Empathy	70.53 (21.22)	1.98	77.22 (34.82)	2.32	60.30 (39.50)	1.70	7.85	.020*	Intentional vs. Accidental	.118	.353
										Intentional vs. Neutral	.219	.656
										Accidental vs. Neutral	2.80	.016*
	Empathic Concern	5.36 (2.97)	2.71	2.48 (3.93)	2.29	−4.36 (3.46)	1.00	64.83	< .001*	Intentional vs. Accidental	1.90	.172
										Intentional vs. Neutral	7.66	< .001*
										Accidental vs. Neutral	5.76	< .001*
	Personal Distress	5.60(3.33)	2.74	1.36(4.01)	2.18	−4.29 (3.82)	1.09	57.36	< .001*	Intentional vs. Accidental	2.52	.036
										Intentional vs. Neutral	7.38	< .001*
										Accidental vs. Neutral	4.86	< .001*
	Moral Evaluation	5.18(3.61)	2.82	−0.79 (4.54)	2.06	−4.79 (3.72)	1.11	61.58	< .001*	Intentional vs. Accidental	3.41	.002*
										Intentional vs. Neutral	7.66	< .001*
										Accidental vs. Neutral	4.25	< .001*
Second Assessment $n = 20$	Cognitive Empathy	72.27 (13.75)	1.85	88.72 (16.99)	2.45	63.25 (32.29)	1.70	6.30	.043*	Intentional vs. Accidental	−1.90	.173
										Intentional vs. Neutral	.474	1.00
										Accidental vs. Neutral	2.37	.053
	Empathic Concern	6.33 (1.71)	2.98	0.54 (3.94)	1.98	−6.85 (2.22)	1.05	37.54	< .001*	Intentional vs. Accidental	3.16	.005*
										Intentional vs. Neutral	6.09	< .001*
										Accidental vs. Neutral	2.93	.010*
	Personal Distress	6.55 (1.93)	2.98	0.27 (3.54)	1.98	−6.54 (2.70)	1.05	37.54	< .001*	Intentional vs. Accidental	3.16	.005*
										Intentional vs. Neutral	6.09	< .001*
										Accidental vs. Neutral	2.93	.010*
	Moral Evaluation	5.42 (3.72)	3.00	−3.01(2.77)	1.95	−6.84 (2.34)	1.05	38.10	< .001*	Intentional vs. Accidental	3.32	.003*
										Intentional vs. Neutral	6.17	< .001*
										Accidental vs. Neutral	2.85	.013*

situational empathy task, but only for three of the 14 variables. Moreover, two of these changes were negative: in one case, the dance group had significantly lower scores for empathic concern for accidental situations, and in the other, the music group had significantly lower scores for personal distress for neutral ones. The most important exception to this trend was that the music group showed a significant increase in their empathic concern responses to intentional harm stimuli in the second assessment. Although this might be the most informative section of the situational empathy measure for our purpose, because it involves sharing the pain of a person who has been victim of a physical attack, the scores of the music group were not significantly higher than those of the dance group when assessed for the second time.

Even if we assume that the absence of significant differences was due to lack of statistical power attributable to the small sample size (the observed statistical power for the Wilcoxon tests that contrasted pre- and post-scores ranged from 0.03 to 0.72), we should at least be able to observe a coherent pattern of results. This is not the case. First, as can be seen in Figure 1 and Table 3, the results from the TEQ suggest that both groups had lower scores on this measure after one year. Second, the music group had slightly lower scores in the prosocial behaviour task in the second assessment (see Figure 2 and Table 3). Third, the positive changes we observed in the intentional harm situations of the empathy for pain task in the second assessment are not reflected in the scores for the accidental harm situations, where both groups tended to have lower scores in the second assessment (see Figures 4–10). Fourth, the fact that, by the time of the second assessment the differences between the dance and the music groups were no longer significant, and that the scores from the dance group stayed the same, suggest that the music group improved after a year of training. However, it is important to note that these positive changes were never large enough for the scores of the music group to exceed those of the dance group.

Finally, two findings may tempt us to conclude that the participants in the music group had better empathy attitudes after one year of training: their improvement in the empathic concern scores for intentional harm situations, and the fact that they caught up with the dance group in most of the variables of the empathy for pain task. However, again, because these findings are not consistent with the other measures, we prefer to interpret them more parsimoniously, as corresponding to the phenomenon of regression to the mean (Barnett et al., 2005).

Unfortunately, the lack of a control group makes it difficult to establish if the changes observed in our participants' scores can be attributed to mere maturation processes, the regression-to-the-mean phenomenon or the effect of the cultural training they underwent. However, a comparison with the findings of a previous study by Gonzalez-Gadea et al. (2014, p. 5), which used the empathy for pain task with a sample of 30 typically developing Colombian participants ($M_{age} = 16$, $SD = .063$), shows that our participants displayed similar performance in both assessments. The scores from both of our groups were less than one standard deviation away from the scores of the Gonzalez and colleagues' sample in the empathic concern, personal distress and moral evaluation scores (mean difference = −2.21). The exception was the cognitive empathy scores in intentional harm situations, where our participants had lower scores in both assessments than their sample (mean difference in first assessment = −26.36; mean difference in second assessment = −22.27).

Study 2: Interview study

The comparative field study found no evidence of improvements in prosociality and empathy among participants in the intervention programmes. However, although these two constructs feature highly in psychosocial research into arts interventions themselves, they do not necessarily capture the full range of constructs relevant to practitioners. In addition to testing for

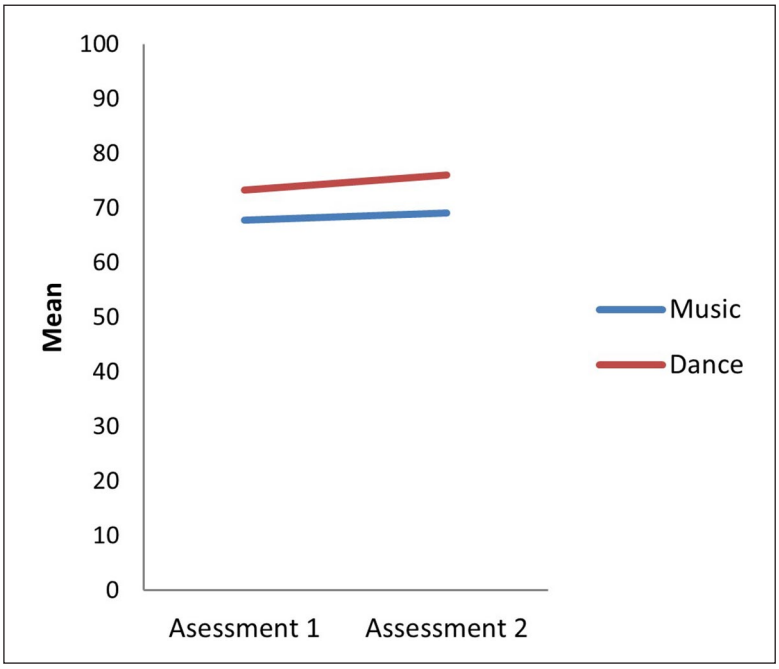


Figure 3. Mean scores: Cognitive Empathy for Intentional Harm Situations

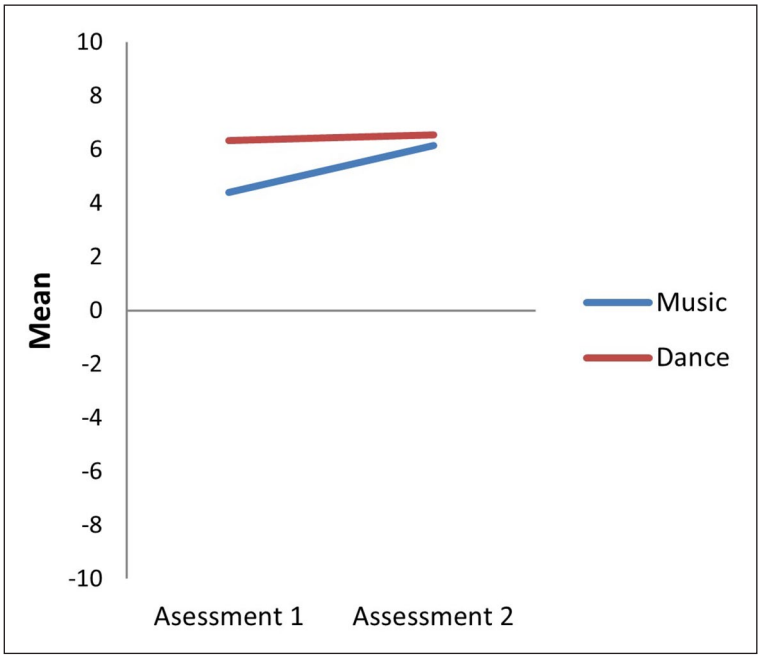


Figure 4. Mean scores: Empathic Concern for Intentional Harm Situations

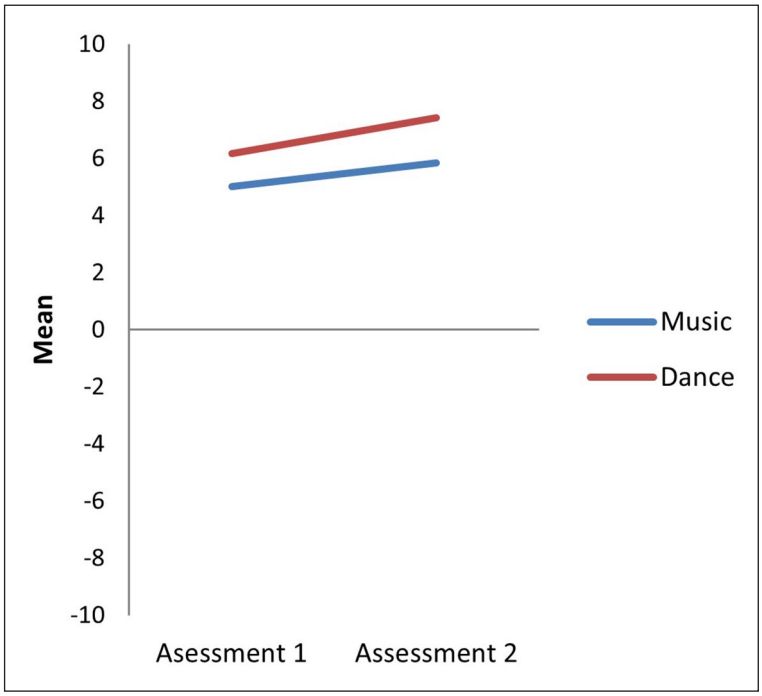


Figure 5. Mean scores: Personal Distress for Intentional Harm Situations

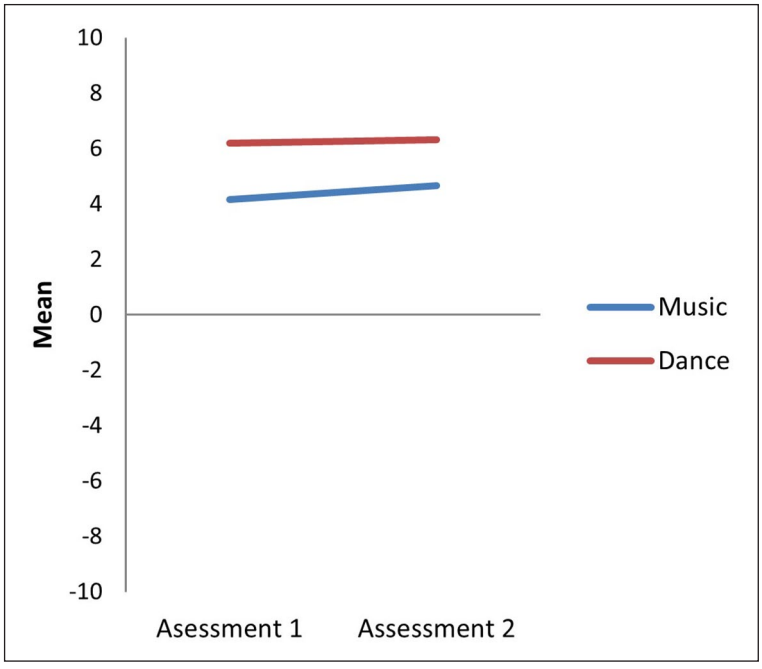


Figure 6. Mean scores: Moral Evaluation for Intentional Harm Situations

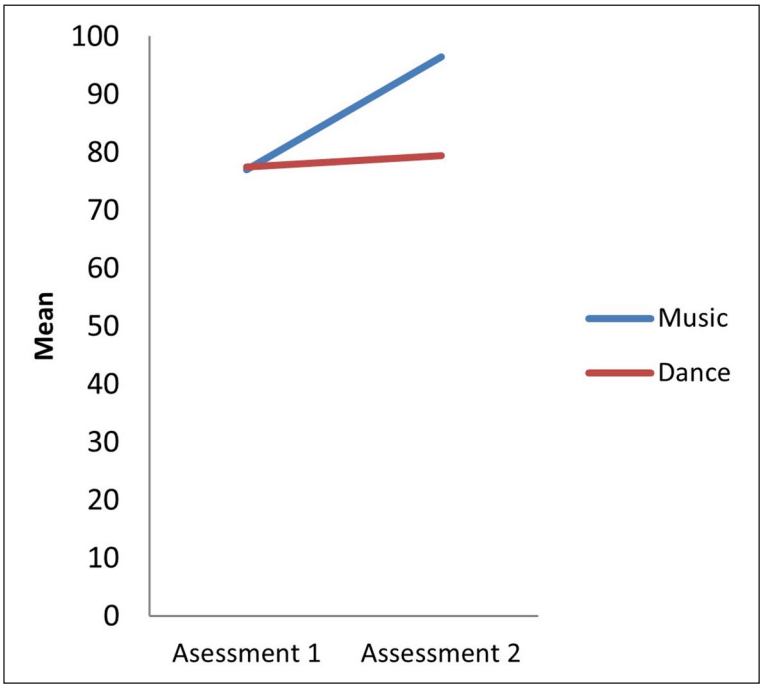


Figure 7. Mean scores: Cognitive Empathy for Accidental Harm Situations

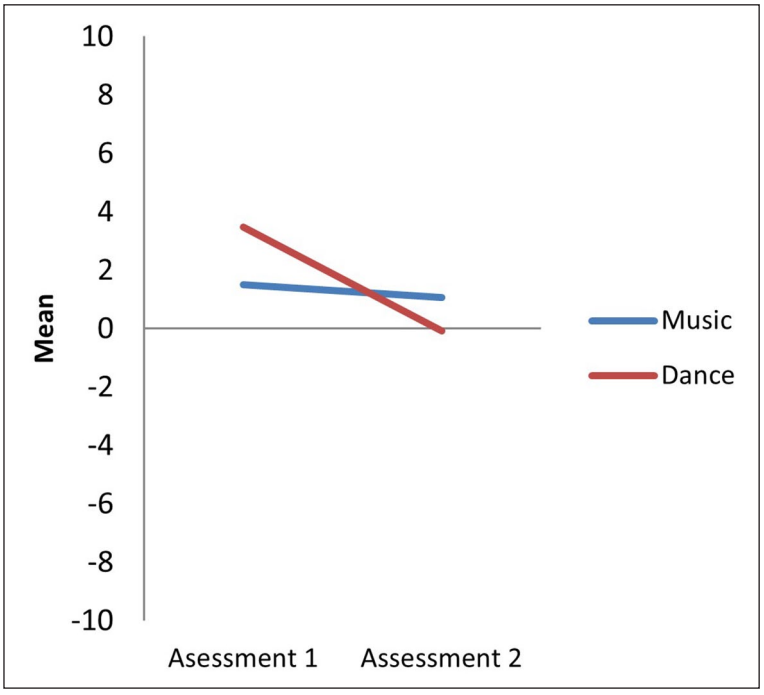


Figure 8. Mean scores: Empathic Concern for Accidental Harm Situations

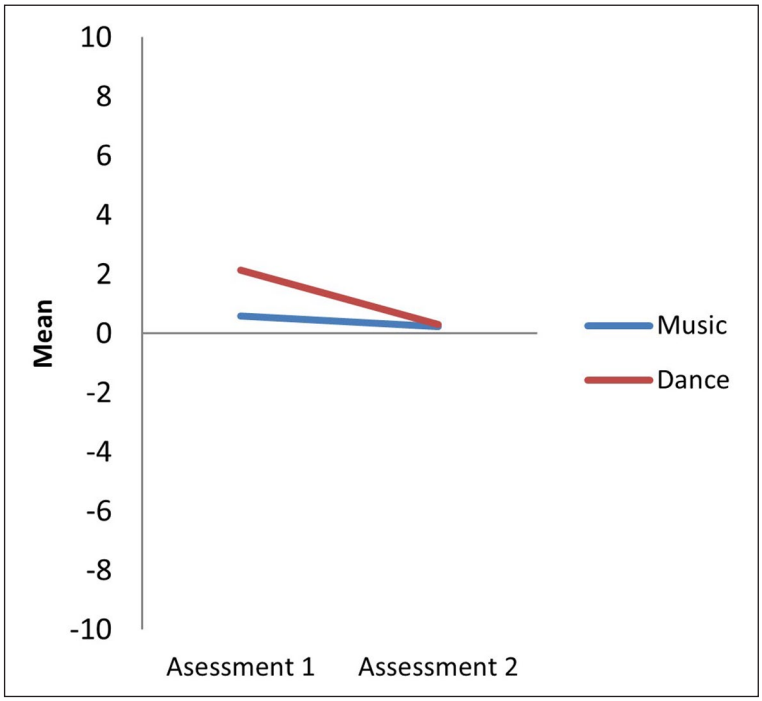


Figure 9. Mean scores: Personal Distress for Accidental Harm Situations

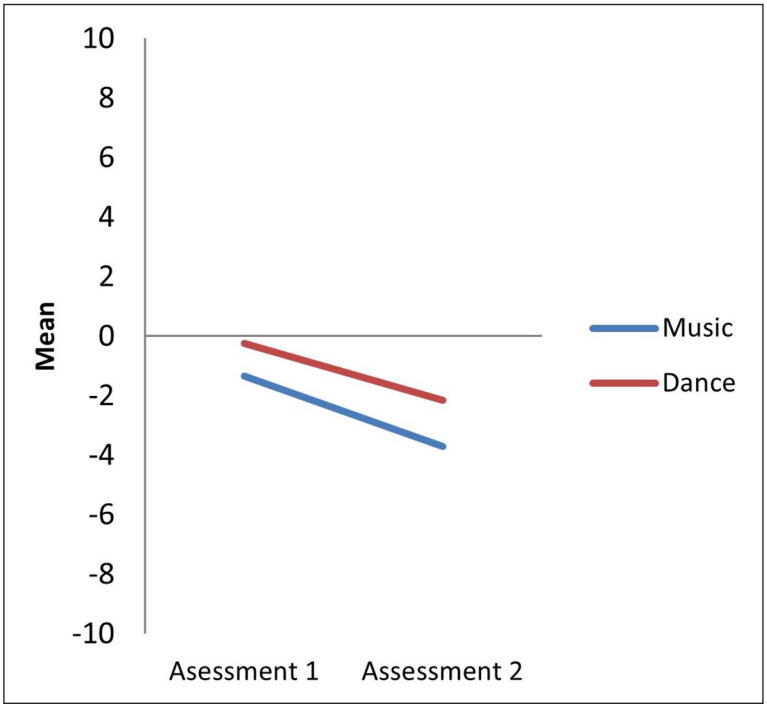


Figure 10. Mean scores: Moral Evaluation for Accidental Harm Situations

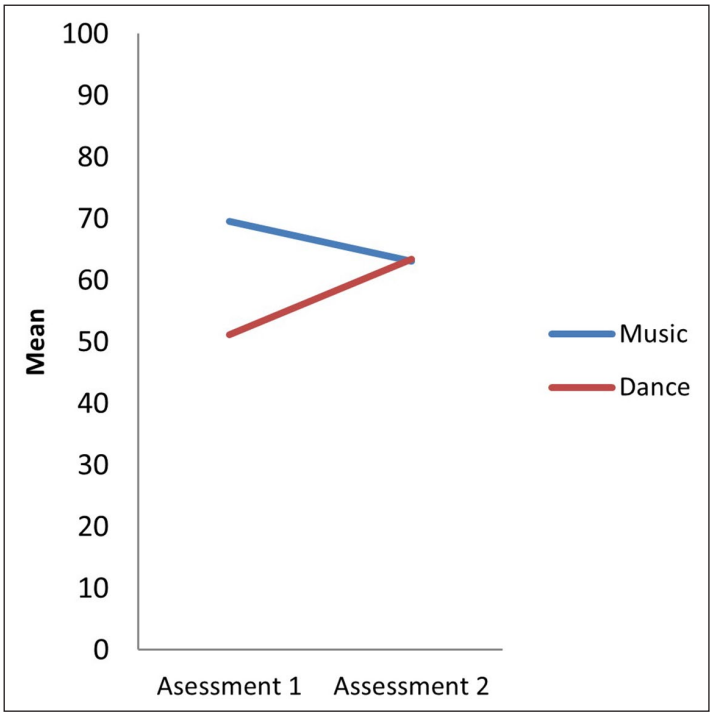


Figure 11. Mean scores: Cognitive Empathy for Neutral Situations

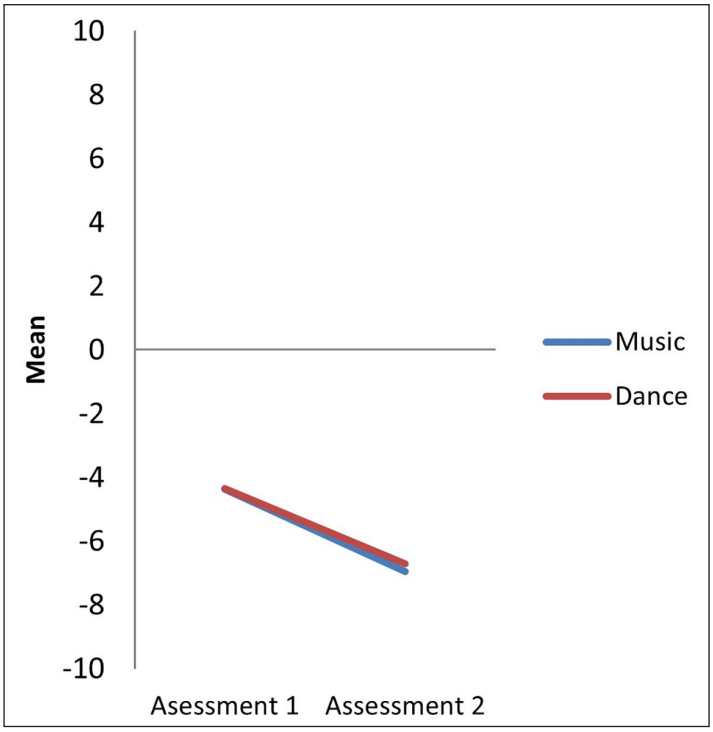


Figure 12. Mean scores: Empathic Concern for Neutral Situations

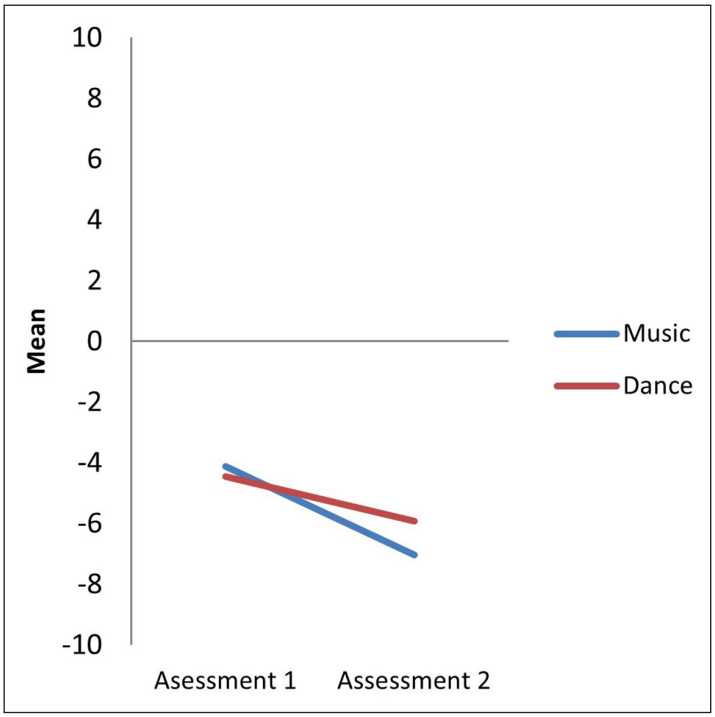


Figure 13. Mean scores: Personal Distress for Neutral Situations

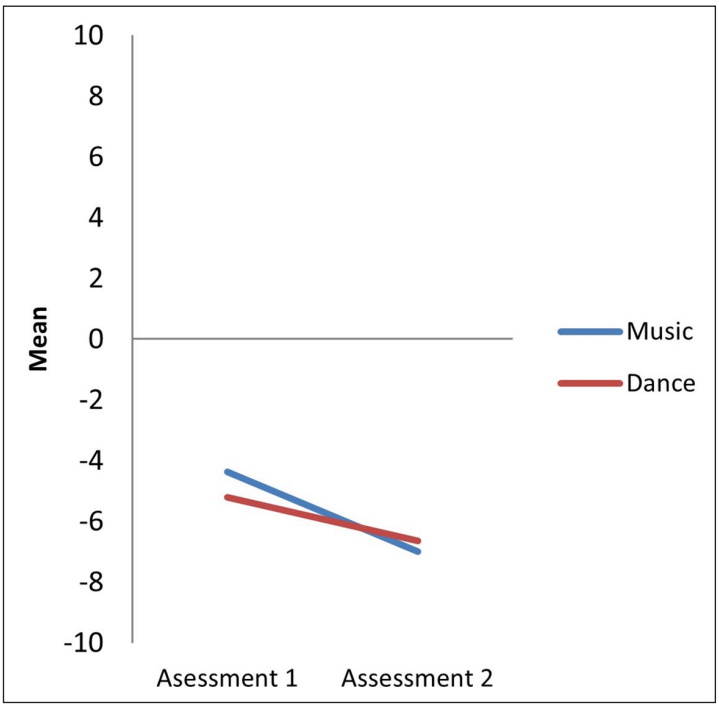


Figure 14. Mean scores: Moral Evaluation for Neutral Situations

longitudinal changes in prosociality and empathy (Study 1), we therefore decided to investigate the intended social impacts of cultural programmes in Cali, including the extent to which prosociality and empathy feature in the intended outcomes (Study 2). We adopt a critical realist perspective in which interviews give us insight into the beliefs and experiences of our participants (Braun & Clarke, 2006).

Method

Participants. Nine participants were recruited from organisations involved in policymaking, funding and/or delivery of cultural programmes for social impacts in the Cali region: three regional policymakers and funders and six from NGOs involved in Study 1. Pseudonyms are used to preserve anonymity.

Data collection and analysis. A semi-structured interview comprising open questions was used to elicit information about the organisation (aims, history, size, governance, structure, funding source, beneficiaries, challenges and aspirations), its practices (type and range of cultural activities offered, impacts, methods of evaluation) and values as experienced by the participant. Telephone interviews of around 45 minutes duration were conducted in Spanish, audio-recorded and transcribed prior to English translation and thematic analysis (Braun & Clarke, 2006).[4]

Findings and discussion

Thematic analysis identified four linked topics regarding the intended social impacts of these cultural intervention programmes: the purpose was to achieve individual and social transformation, evidenced as positive outcomes, by creating the conditions for transformation, aspects of which were unique to different cultural activities. These are elaborated on in the following.

Informants described their organisations as working with young people living in contexts of high levels of deprivation and criminality to transform each individual's skills and life opportunities:

> The people who are part of the projects develop skills: our approach is the development of life skills. Through the projects they develop social, emotional, and cognitive skills that allow them to understand and transform their realities and the possibility of building life projects. In this case . . . it has to do with the issue of violence, the issue of (drugs) micro-trafficking, with the issue of the culture of illegality. (Director, industry-sponsored youth symphony orchestra)

Intervention programmes were framed as helping the individual to develop socio-emotional and cognitive skills. Anecdotal evidence of success referenced individual outcomes: cognitive effects of music on memory and reading from instrumental learning; psychosocial and

4 In the familiarisation phases, the Spanish transcription was read repeatedly by the first author and large chunks of the text were identified for further analysis. These sections were then translated by the authors and familiarisation notes compared. We approached the data inductively, identifying patterns in the data to form codes and themes, and deductively drawing on psychological theoretical constructs of group music making, prosociality and empathy. The coding began with semantic codes, and latent codes were identified as we immersed ourselves in the data. Themes were formed through an iterative process of grouping codes, and recording until both authors were satisfied that the set of themes reflected the data set as a whole.

emotional effects on empathy, social interaction and improved emotion regulation, including "feeling the other, of respecting the other, of creating coexistence for the other" (Director, infant and youth symphony orchestra); and employability and access to further training. The construct of the *life project* was used by many informants to describe the ability of individuals to envisage a (better) life, enabled by these skills.

A second construct of *social transformation* referred to broader societal changes aimed at achieving social inclusion:

> The ultimate goal is to give people an opportunity to transform themselves individually and that they provide a service to the community, that is, that they become aware that they are citizens and that they can actively participate as conscious citizens in the transformation of its environment and its territory. (Director, arts organisation)

This account construes individual transformation as development of personal capacities including the "ability to interact with the community," and as a means to inclusion and participation as citizens that will enable them to change their surroundings.

The intentions of these organisations can be characterised as creating the conditions for individual and social transformation. This manifests in explicit social work with children and their families, and annual themes (such as *culture of peace*) worked on in parallel with the cultural activity, and implicit forms of desired behaviours, including required school attendance, discipline and a regular routine and substantial time commitment. Routine, discipline and reducing time available to participate in harmful opportunities were seen as conditions for individual life transformation:

> It is using music as a tool to transform the lives of the beneficiaries or students who are part of the foundation. It is using music so that they have an activity that occupies their free time and that keeps them off the streets, out of gangs, out of violence. (Director, orchestral music-training programme)

Collective behaviour was construed as a means for *transformation*. It was viewed as enabling individuals to connect with each other: as a product of the amount of time spent together on a meaningful activity; as a "meeting space" (Director, industry-sponsored youth symphony orchestra); as a means of enacting social formation ("an orchestra already implies socialization, a community . . . the orchestra is like a new family, so they learn from that social integration . . . and music also allows them to unite to form one" (Director, infant and youth symphony orchestra); and as an emblem of "social inclusion, rights and diversity" (Regional Director, national orchestra school).

Musical excellence has a complex role in the context of these programmes, not least because one way of achieving social inclusivity is to offer mass music-training programmes with perceived limitations on the achievement of musical excellence. Paradoxically, even though informants attributed a limited role in transformation to musical excellence (with the exception of affording a few participants a musical livelihood, and acting as a cultural signifier) they illustrated the success of their programmes by reference to accounts of individuals who had gained places on higher education music programmes or pursued professional performing careers.

Different cultural activities were viewed by some as affording different conditions for transformation:

> Regarding violence prevention and community transformation, the children in the football programme are more concerned with transforming the dynamics of violence in their territory; those of the orchestra are more oriented to their own life project. We have tried to balance these things, because

> what we observe is that the orchestra children develop these skills very well and their life projects are very reflective and critical, but when it comes to thinking about their territory and influencing their territory to transform it, it's not that strong. (Director, industry-sponsored youth symphony orchestra)

For this informant, the music-training project fosters conditions for individual transformation, but this does not translate into an orientation towards change in their home community compared to the football project. This highlights useful avenues for future research into the way that different activities afford specific outcomes. In this case, one possibility is that, because the football projects recruit children from single, local communities, and involve competitive neighbourhood football matches, it lends itself to territorial identification. By contrast, children in the orchestra, run by the same NGO, are drawn from and perform in several neighbourhoods, which may create less sense of affiliation to a single community beyond the orchestra. Hence, while different cultural programmes might share common intentions regarding individual and social transformation, and commonalities in how it is afforded, they may be unequal in their ability to achieve these goals.

Analysis of the discourse on the relationship between the cultural activity and the community offers deeper insight as to why this might be. Informants from the music-training projects described the individual transformation, whether actual or intended, as a journey away from the home community, and by implication its associated criminality and poverty, saying, for example, "that music, somehow, helps them escape" (Director, infant and youth symphony orchestra), towards the world of music constructed as a particular orchestra and/or as the classical music world more generally: ". . . it was bringing the children closer to classical music . . . to bring these children from this community closer to the orchestra Philharmonic or classical music" (Director, non-profit music school). Music and the community are constructed as two separate worlds, demarcated by music genres and associated lifestyles:

> [W]hat this type of music does is to open themselves to a different world. So you see the young people who listen to reggaeton, who listen to other types of music, and these children from the foundation, I don't claim that they don't [listen to reggaeton], but they are more immersed in another style of music and in another lifestyle. (Director, orchestra music training programme)

If, as this implies, the two worlds are separate, transformations that take place for the individual stay with that individual rather than benefiting the community from which they come, which is a criticism also made of Venezuela's El Sistema programme (Mota et al., 2016).[5]

This analysis reveals a potential disjunction between the central constructs underlying these cultural intervention programmes and those of psychologically informed research. Informants identified intended individual and social outcomes and a path by which the one might lead to the other, but the evidence used by organisations and funders, including the numbers and demographics of people impacted, and testimonials, and the constructs deemed central by

5 Perhaps the most well-known example of such a programme involving music training is the Venezuelan National System of Youth and Children's Orchestras, El Sistema, aimed at achieving social transformation through engaging poor children, and now the subject of sustained academic critique. This critique notes the lack of evidence for significant impacts from El Sistema, and the mismatch between its methods (grounded in its history as a centre for orchestral training) and its narrative of social transformation (Baker, 2014). The presumption in its narratives is that those in the included centre can provide what is lacking in an excluded periphery, possibly to the extent of obscuring the bigger problem of social inequality (Mota et al., 2016).

psychologically informed research, such as empathy and prosociality, do not reflect the range or relative importance of intended programme outcomes.

General discussion

The aim of our research was to understand the prosocial impacts of third sector cultural activities with children and adolescents in impoverished and violence-stricken urban neighbourhoods in Cali, Colombia. We carried out an exploratory comparative field study that assessed the development over one year of empathic attitudes and prosocial behaviour in children who were enrolled in music, dance or football training social intervention programmes. In addition, we interviewed the directors of cultural intervention programmes and local government policymakers and funders, to understand the place of individualised psychosocial constructs such as prosociality and empathy in the aims and outcomes of such projects. The main finding from our quantitative comparative study was that the children who participated in the dance-training programme had slightly higher scores for cognitive and affective empathy, as measured by the situational empathy task, than the children in the music-training programme. However, one year later, neither the music nor the dance group exhibited significant improvements in any of the measures: trait empathy, prosocial behaviour or situational empathy. The main finding of our qualitative study was that the outcomes intended by those who fund and deliver cultural intervention programmes were neither well measured by the organisations themselves, nor well served by existing psychological constructs of empathy and prosocial behaviour.

One way of interpreting the lack of evidence for impacts of the programmes on prosociality and empathy is to attribute them to the methodological limitations of our study. It is possible that the effects of these programmes on empathic attitudes were not captured by the measurement techniques that we implemented, because they introduced demand characteristics or were too artificial. For instance, although our measure of prosocial behaviour has some advantages over measures that rely on the participants' self-report, this technique may work better in younger children who may not realise the intention behind the experimenter's action of dropping the pencils on the floor (e.g., Kirscher & Tomasello, 2010; but see Kokal et al. 2011 for a study that used this technique with adult participants). In addition, the empathy for pain task can also be considered as too artificial, because it involves watching scenes of people being harmed by another on a computer screen, without providing any additional context for their motives or relationship. However, even with its limitations, this task has some advantages over self-report questionnaires, and over more ecologically valid stimuli such as narrative films, which may introduce more demand characteristics and confounding factors in the participants' empathic responses. Moreover, as mentioned previously, a comparison of results for the empathy for pain task from our sample with a similar one suggests our results are reliable. It is also possible that, since our sample consisted of pre-existing groups of children who had already decided to enrol in these cultural programmes, our sample was biased by the fact that our participants probably came from families who cared about their education and moral development. However, this is not reflected in the results of the empathy for pain task, where again, the scores were similar to those of an equivalent group of Colombian teenagers who had not participated in such cultural activities. Moreover, even if our sample was biased, this does not explain why the groups did not improve in their scores in the second assessment, given that we did not find a ceiling effect in the first assessment. In addition, because we could not assess the children before they started receiving any training due to the recruitment practices of the NGOs (some participants were tested when they had spent up to three months in their programme), then it is plausible that, at least

for some of the children, the effects of the cultural training they underwent happened during the first weeks of their enrolment in their programmes. However, if this is the case, it is still difficult to explain why this sudden improvement in empathic attitudes and prosocial behaviours did not continue to increase during one year of training.

Finally, it is possible that significant improvements could not be observed because it was premature to assess the children after only one year of training. However, previous studies such as those of Rabinowitch et al. (2013) have found improvements in empathy tasks after nine months of music training, and we did not even observe a consistent trend of improvement in the data, which rules out this interpretation.

A second explanation for the results of our comparative study is that the empathy-boosting effects of these cultural training programmes may be limited to in-group members. That is, as found in previous studies (e.g., Cirelli et al., 2014; Kirschner & Tomasello, 2010; Lonsdale & North, 2009; Tarr et al., 2015), it is possible that participating in these cultural activities only makes it more likely that children will feel empathy and display altruistic behaviours towards the people who were present during their training situations, such as classmates, teachers, coaches, etc. However, other research has found that participating in cross-cultural music listening (Sousa et al., 2005), musical rhythmic games (Tunçgenç & Cohen, 2016) and other artistic activities (Van de Vyver et al., 2019) increase altruistic behaviours towards out-group members.

Alternatively, a third way of interpreting our results is to conclude that, contrary to our expectations, participating in these cultural activities does not improve empathic attitudes and prosocial behaviours, neither in the short- nor the mid-term.

Does this mean that these programmes do not offer any long-term psychosocial benefits to these children? Such a pessimistic conclusion is probably unwarranted. The results of our qualitative study suggests that, whereas music psychologists have characterised the effects of musical training in terms of empathy and prosocial attitudes, the impact of these interventions resides in their ability to help children envision a positive life project. This suggests that, as a community of scholars, we may need to measure different variables: self-esteem, self-efficacy beliefs, life-project building, and social identity. The role of music instrument learning as a means to a professional career arises as one potential outcome, but for the programmes we investigated the main role of instrumental learning offered to all, regardless of ability, is instead its ability to create conditions for the life project. This is very different from El Sistema, for instance, whose focus has been much more on selection for and production of orchestral players (Baker, 2014). Moreover, many of the programmes we investigated explicitly support psychosocial development through provision of social work with the participants rather than leaving it to arise spontaneously from cultural participation.

Other aspect of these programmes highlighted by the qualitative study was the way in which certain kinds of activity were more or less well positioned to enable the development of social as well as individual outcomes. The recruitment structure and competitive activities of the football project were perceived as lending themselves to the development of territorial identity to a much greater extent than the orchestral project. At the very least, even if the outcomes from these programmes are not apparent from quantitative psychological measures, given the problematic contexts in which these children live, neighbourhoods ruled by gang violence, drug trafficking and consumption and high rates of school dropout, these programmes provide children with spaces in which they can use their spare time in positive activities. In other words, participating in these programmes makes these children less likely to engage in antisocial activities, simply because they leave them with little spare time to do so. Nevertheless, this is against a context in which such programmes, whatever their good intentions may fail to value chil-

dren's own home cultures and superimpose a culture from the included centre thereby re-enacting inequality.

Challenges and implications

This study highlights some of the practical and conceptual challenges of working in this field. Researching pre-existing cultural intervention programmes involves practical difficulties: the ethical dimensions of working with children, their families and the charitable foundations; the difficulty of a longitudinal study with a changing population; differences in the recruitment to and frequency of participation in different programmes making direct quantitative comparison inappropriate. The conceptual difficulties are also substantial. Our analysis indicates that the constructs researchers bring to such studies may not be those best suited to the intended outcomes of the programmes. Moreover, the psychosocial impact seems evident to practitioners who feel confident that their practice is effective in producing benefits[6] and do not see the need to have it further evidenced. There is an assumption that "music does good," such that neither practitioners nor sponsors necessarily see a need for better evidence of the social impacts of music-training programmes.

There was a particular assumption that classical music is inherently good, and has better effects than other musical repertoires. It is safe to assume that music-training programmes can have positive effects on building skills such as discipline, perseverance and cooperation. However, we question whether the traditional pedagogical structure of classical music, with its rigid vertical hierarchies, is optimally suited to afford the construction of democratic relations among the children who participate in this kind of training, as critiqued in Geoff Baker's (2014) analysis of El Sistema. These same qualities might well be developed in learning of other musical skills that could offer more immediate employability outcomes, such as popular music and music production. However, to do this would challenge deeply held beliefs about the association between certain repertoires and certain lifestyles. Finally, it is notable that there is limited focus on the effects of these interventions beyond the individual: how do these programmes impact their families, neighbourhoods, cities and the nation? And to what extent can these impacts be captured by psychological instruments that focus on individual, short-term effects?

Accepting the limitations of our methods, we suggest on the basis of our results that music training is not associated with larger effects on the development of prosocial attitudes than dance training. Consequently, researchers and practitioners need to be more sceptical and reflective about their assumptions in this regard. Moreover, we would question the extent to which constructs of prosociality and empathy are indeed the right focus for research in this domain, when cultural intervention programmes themselves champion a wider set of goals related to individual and social transformation. This highlights the need to work with these nonacademic partners in critically reflective ways to develop research programmes together.

Acknowledgements

We would like to thank Ana Cocuy, Valentina Plazas, Andrey Chavez, Kevin Cuellar, Laura Sanchez, Santiago Murillo and Steven Vargas for their help in collecting the data for the studies reported here.

6 When asked what they would do with more funding (Study 2), organisations' aspirations were for further reach (increasing the number of young people on programmes) rather than deepening the significance of the impacts through their activities, implicitly suggesting satisfaction with this aspect of their programmes.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by an award from the Academy of Medical Sciences Global Challenges Research Funding Networking (grant number GCRFNGR2\10079).

ORCID iD

Julian Cespedes-Guevara https://orcid.org/0000-0002-8816-3650

References

Auné, S. E., Abal, F. J. P., & Attorresi, H. F. (2016). Diseño y construcción de una escala de conducta prosocial para adultos. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, *42*(2), 15–25. https://doi.org/10.21865/ridep42_15

Baker, G. (2014). *El Sistema: Orchestrating Venezuela's youth*. Oxford University Press.

Baker, G., Bull, A., & Taylor, M. (2018). Who watches the watchmen? Evaluating evaluations of El Sistema. *British Journal of Music Education*, 1–15. https://doi.org/10.1017/s0265051718000086

Barnett, A. G., Van Der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, *34*(1), 215–220. https://doi.org/10.1093/ije/dyh299

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Broadwood, J., Bunting, C., Andrews, T., Abrams, D., & Van de Vyver, J. (2012). *Arts & kindness*. People United Publishing.

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, *17*(6), 1003–1011. https://doi.org/10.1111/desc.12193

Decety, J., Michalska, K. J., & Kinzler, K. D. (2012). The contribution of emotion and cognition to moral sensitivity: A neurodevelopmental study. *Cerebral Cortex*, *22*(1), 209–220. https://doi.org/10.1093/cercor/bhr111

DNP Columbia (Departamento Nacional de Planeación de Colombia) (2015, 16 June). *Plan Nacional de Desarrollo 2014-2018*. Presidencia de la República de Colombia. http://wp.presidencia.gov.co/sitios/especiales/Documents/20150616-especial-plan-nacional-desarrollo/index.html

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*(1), 143–180. https://doi.org/10.1111/j.1751-2409.2010.01020.x

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th Edition). SAGE.

Goldstein, T. R., & Winner, E. (2012). Enhancing empathy and theory of mind. *Journal of Cognition and Development*, *13*(1), 19–37. https://doi.org/10.1080/15248372.2011.573514

Gonzalez-Gadea, M. L., Herrera, E., Parra, M., Gomez Mendez, P., Baez, S., Manes, F., & Ibanez, A. (2014). Emotion recognition and cognitive empathy deficits in adolescent offenders revealed by context-sensitive tasks. *Frontiers in Human Neuroscience*, *8*, 850. https://doi.org/10.3389/fnhum.2014.00850

Kawase, S., Ogawa, J., Obata, S., & Hirano, T. (2018). An investigation into the relationship between onset age of musical lessons and levels of sociability in childhood. *Frontiers in Psychology*, *9*, 2244. https://doi.org/10.3389/fpsyg.2018.02244

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354–364. https://doi.org/10.1016/j.evolhumbehav.2010.04.004

Kokal, I., Engel, A., Kirschner, S., & Keysers, C. (2011). Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment-if the rhythm comes easily. *PLoS One*, *6*(11), e27272. https://doi.org/10.1371/journal.pone.0027272

Lamont, A., & Ranaweera, N.A. (2019). Knit one, play one: Comparing the effects of amateur knitting and amateur music participation on happiness and wellbeing. *Applied Research in Quality of Life*, 1–20. https://doi.org/10.1007/s11482-019-09734-z

Lonsdale, A. J., & North, A. C. (2009). Musical taste and ingroup favouritism. *Group Processes & Intergroup Relations*, *12*(3), 319–327. https://doi.org/10.1177/1368430209102842

Mota, G., Baker, G., Ilari, B., O'Neill, S., & Senyshyn, Y. (2016). Social inclusion and non-formal music learning—in the shadow of Venezuela's *El Sistema*. In R. Wright, B.A. Younker, & C. Beynon (Eds.), *21st Century Music Education: Informal Learning and Non-Formal Teaching*, *7* (pp. 38–54). Canadian Music Educators' Association.

Pearce, E., Launay, J., & Dunbar, R. I (2015). The ice-breaker effect: Singing mediates fast social bonding. *Royal Society Open Science*, *2*(10), 150221. https://doi.org/10.1098/rsos.150221

Raabe, T., & Beelmann, A. (2011). Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of age differences. *Child Development*, *82*(6), 1715–1737. https://doi.org/10.1111/j.1467-8624.2011.01668.x.

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*(4), 484–498. https://doi.org/10.1177/0305735612440609

Robertson, C., Tarr, B., Kempnich, M., & Dunbar, R. (2017). Rapid partner switching may facilitate increased broadcast group size in dance compared with conversation groups. *Ethology*, *123*(10), 736–747. https://doi.org/10.1111/eth.12643

Schellenberg, E. G., Corrigall, K. A., Dys, S. P., & Malti, T. (2015). Group music training and children's prosocial skills. *PLoS One*, *10*(10), e0141449. https://doi.org/10.1371/journal.pone.0141449

Sloboda, J.A. (2019, 26–27 July). *The SIMM research landscape: An analysis of research presented at SIMM events in 2017, 2018, 2019* [Spoken presentation]. 4th SIMMposium, Bogotá, Colombia. http://www.simm-platform.eu/wp-content/uploads/2019/07/Sloboda.pdf

Sousa, M. D. R., Neto, F., & Mullet, E. (2005). Can music change ethnic attitudes among children? *Psychology of Music*, *33*(3), 304–316. https://doi.org/10.1177/0305735605053735

Spreng, R. N., McKinnon, M. C., Mar, R. A., & Levine, B. (2009). The Toronto Empathy Questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of Personality Assessment*, *91*(1), 62–71. https://doi.org/10.1080/00223890802484381

Tarr, B., Launay, J., Cohen, E., & Dunbar, R. (2015). Synchrony and exertion during dance independently raise pain threshold and encourage social bonding. *Biology Letters*, *11*(10), 20150767. https://doi.org/10.1098/rsbl.2015.0767

Tay, L., Pawelski, J. O., & Keith, M. G. (2018) The role of the arts and humanities in human flourishing: A conceptual model. *The Journal of Positive Psychology*, *13*(3), 215–225. https://doi.org/10.1080/17439760.2017.1279207

Tunçgenç, B., & Cohen, E. (2016). Movement synchrony forges social bonds across group divides. *Frontiers in Psychology*, *7*, 782. https://doi.org/10.3389/fpsyg.2016.00782

Van de Vyver, J., & Abrams, D. (2018). The arts as a catalyst for human prosociality and cooperation. *Social Psychological and Personality Science*, *9*(6), 664–674. https://doi.org/10.1177/1948550617720275

Van de Vyver, J., Abrams, D., Spinner, L., Pelletier, J., Ali, S. Y., & Kapantai, I. (2019). Participatory arts interventions promote interpersonal and intergroup prosocial intentions in middle childhood. *Journal of Applied Developmental Psychology*, *65*, 101069. https://doi.org/10.1016/j.appdev.2019.101069

Williams, E., Dingle, G. A., Jetten, J., & Rowan, C. (2019). Identification with arts-based groups improves mental wellbeing in adults with chronic mental health conditions. *Journal of Applied Social Psychology*, *49*(1), 15–26. https://doi.org/10.31234/osf.io/zhm7u