

Building harmony: Reducing and measuring racism in Australian schools

Kaine Grigg
 Monash University
 Lenore Manderson

University of Witwatersrand and Monash University

Racism is a pervasive and persistent social problem that has widespread negative effects on society, with physiological and psychological impacts extending to children. Various school-based anti-racism and pro-diversity initiatives have been implemented throughout Australia, but few have been efficacious. Moreover, until recently such programs were unable to be evaluated appropriately due to lack of suitable instruments. RACES was developed as a measure of racism, acceptance, and cultural-ethnocentrism to evaluate the impact of such interventions, in the first instance a project known as Building Harmony. The research reported in this article investigated the reliability and validity of RACES, while exploring the efficacy of Building Harmony in reducing racism and social, emotional, and behavioural difficulties. Participants were 296 students enrolled in years five or six at six primary schools in a growth corridor in the southeastern region of Melbourne, Australia. RACES, MCSDS-A, and SDQ were completed to enable evaluation of the psychometric properties of RACES (internal consistency, test-retest reliability, convergent validity, discriminant validity, and predictive validity) and to assess the impact of Building Harmony. Results provide strong support for RACES as a reliable and valid measure of racist attitudes in the Australian context. Although only tenuous evidence for the efficacy of Building Harmony was found, ongoing evaluation of similar initiatives will ensure that more focused and effective racism-reduction interventions can be implemented. This will support reductions in community levels of racism and the enhancement of acceptance of diversity within Australian schools.

Child-specific research has documented the psychological and physical effects of racism. These include anxiety, depression, and hopelessness; increased alcohol, tobacco, and drug use; lowered self-esteem, self-worth, self-efficacy, and life satisfaction; anger, conduct problems, delinquency, and Attention Deficit and Hyperactivity Disorder; and indicators of metabolic and cardiovascular disease (Howarth, 2009; Pachter & Coll, 2009; Priest et al., 2012). Ethnographic research in primary and secondary schools in a number of Australian states has demonstrated how students, parents, and teachers support and perpetuate widely prevalent racist behaviour and attitudes (Ben-Moshe & Halafoff, 2014; Gross & Rutland, 2014).

Given the potentially devastating impacts of racism on children and its prevalence in educational institutions, the implementation of school-based interventions is considered integral to reduce racism and its effects (Greco, Priest, & Paradies, 2010). A range of school-based and child and

adolescent focussed initiatives aimed at reducing racism and prejudice more broadly have been implemented worldwide, although primarily in the US (Beelmann & Heinemann, 2014; Verkuyten & Thijs, 2013). A recent meta-analysis of child and adolescent prejudice reduction interventions demonstrated that diverse approaches have been utilised, with low to moderate effects (Beelmann & Heinemann, 2014). A range of moderators on program effectiveness were highlighted. Programs based on direct intergroup contact experiences or social-cognitive training and those involving majority group participants showed the strongest effects. Corresponding positive findings with imagined intergroup contact in British children have also been found (Stathi, Cameron, Hartley, & Bradford, 2014). These analyses demonstrate that structured interventions can be effective in children and adolescents, but the distinctive nature of the assessed initiatives means that their translation to different settings is challenging, if not unachievable.

Research focusing on early childhood interventions has been less encouraging, partly because of inherent socio-cognitive limitations, which can lead young children to resist anti-racism messages from external communicators and educators (Johnson & Aboud, 2013). A recent systematic review of interventions to reduce prejudice and enhance inclusion and respect for ethnic differences in early childhood found overall equivocal evidence for their effectiveness (Aboud et al., 2012). Of all effects, only 40% were positive, with attitude effects the most positive. Method of delivery was important: programs using audio-visual or print material proved more positive than interventions based solely on intergroup contact. Consistent with evaluations extending into adolescence, majority group individuals experienced the most positive effects from interventions. In addition, explicitly theory-driven interventions appeared to have the strongest results, highlighting the need for robust research designs and the systematic and explicit use of theories of change in future interventions.

The international research community has naturally informed practice in Australia. To date, despite various interventions being implemented in Australian schools, few have been strongly supported empirically (Greco et al., 2010; Walker & Crogan, 1998). Notable exceptions include one study with Muslim and Christian secondary school students in Sydney, which demonstrated affective intergroup bias reduction maintained after 12 months post-intervention (White, Abu-Rayya, & Weitzel, 2014), although only for Muslim and not for Christian students. The Enhancing Relationships in School Communities project implemented in Melbourne primary schools aimed to impact upon and enhance overarching school culture, with qualitatively positive effects on opposition to racism, respecting and supporting diversity, general cultural awareness, and enhancing management of intergroup conflict (Wertheim, Davis, Freeman, & Trinder, 2010). The School Support Programme, operating in primary and secondary schools

in Victoria, was similarly effective in promoting social inclusion and improved outcomes for refugee students and families (Block, Cross, Riggs, & Gibbs, 2014). The Building Racial Harmony component of the overarching Exploring Race and Ethnicity project, implemented in secondary schools in Queensland, was also successful in addressing and amending problematic views of race, ethnicity, culture, and religion (Hickey & Austin, 2009). An additional project yet to be evaluated is Building Harmony in the Growth Corridor (henceforth Building Harmony), an anti-racism and pro-diversity initiative implemented in local primary schools each year by a non-government organisation, Windermere Child and Family Services, and the local government, Cardinia Shire Council (Grigg & Manderson, 2013). The strategy was introduced in 2009 to build the capacity of young people from diverse backgrounds, and their families, to live harmoniously in one of Victoria's emerging communities. Its goal was to discourage crime based on race and to encourage acceptance of diversity.

Many anti-racism, pro-diversity, and cultural awareness projects are not designed to specifically target racism. Although such initiatives often have positive and notable outcomes, demonstrated by both qualitative and quantitative data, they may not necessarily have a meaningful and lasting impact on the level of racism and racist attitudes within the participating schools. Part of the reason for the limited efficacy evidence is that, until recently, there was no appropriate tool to assess the impact of multi-group racism-reduction and pro-diversity strategies (Grigg & Manderson, 2014a). Hence, the impact on the level of racist attitudes of participants in various programs, projects, and initiatives could not be adequately evaluated.

International research has demonstrated that reliable and valid instruments for measuring the levels of prejudice and racism in children and youth can be developed and widely disseminated (Aboud & Doyle, 1996; Tredoux, Noor, & de Paulo, 2009). In Australia, scales have

been developed to assess racist attitudes, but these either concentrate on a specific group (e.g., Indigenous Australians; Pedersen, Beven, Walker, & Griffiths, 2004), or do not have peer reviewed research detailing their development and validation (e.g., Dunn & Geeraert, 2003). For youth, the available instruments are limited to measures of social distance and stereotyping (e.g., Doyle & Aboud, 1995; Walker & Crogan, 1998), instruments that have been adapted from non-Australian measures, but have not subsequently undergone further validation (e.g., White & Gleitzman, 2006), or complex surveys requiring extrapolation of participant prejudice levels from their interpretation of the seriousness of prejudice-laden scenarios, leading to questions of reliability and validity (e.g., White & Abu-Raya, 2012). Therefore, there is a dearth of instruments for accurately measuring racism across groups in Australian youth.

As there is stronger validity in utilising multiple methods (Mullan, Todd, Chatzisarantis, & Hagger, 2014; Tuffin, 2008), both qualitative and quantitative research is required to provide a complete understanding of racism and its impacts. Qualitative methods provide data that can deliver valuable insights into racism as a construct, its development, and its experience. Quantitative methods establish the prevalence of racism, differences among groups, and its relation to other concepts. Qualitative research is necessary to inform the development of quantitative instruments, which in turn are necessary to establish the effectiveness of intervention programs to reduce racism. Proven effectiveness in turn will enable these programs to be disseminated with the goal of reducing community levels of racism.

Given the limitations of existing tools, the accurate appraisal of multi-group racism-reduction and pro-diversity strategies, anti-racism programs, projects, and initiatives in schools is rare. To redress this, a measure of racial, ethnic, cultural, and religious acceptance – the Australian Racism, Acceptance, and Cultural-Ethnocentrism Scale (RACES; Grigg & Manderson, 2014a)

– was developed with children, adolescents, and adults. RACES was developed specifically for a Victorian school-based pro-diversity and racism reduction project (Building Harmony) to measure the acceptance of diversity among participants prior to and post involvement in the intervention activities; the aim was to provide evidence of the effectiveness of the initiative in improving the race-related attitudes of participating students. This article provides additional evidence that RACES can effectively evaluate the effectiveness of anti-racism and pro-diversity initiatives in Australian schools and reports the efficacy findings of the Building Harmony strategy.

Method

Research Setting

The research setting was a southeastern growth corridor region of Melbourne, specifically Officer, a small town in Cardinia Shire approximately 55km from Melbourne city. The City of Casey and Cardinia Shire, neighbouring localities of the City of Greater Dandenong, are two of the most rapidly growing residential areas of Melbourne and Australia, with population estimates far exceeding mean projected growth forecasts at state and national levels (Australian Bureau of Statistics, 2008; Cardinia Shire Council, 2012; City of Casey Council, 2012). Projected increases are of 102% and 166% for Casey and Cardinia from 2006 to 2031 (Cardinia Shire Council, 2012; City of Casey Council, 2012), in contrast with Victorian and Australian growth forecasts of 30% and 32% respectively from 2006 to 2026 (Australian Bureau of Statistics, 2008). The majority of Cardinia Shire residents are Australian born with both parents also born in Australia, rates that are much higher than the general Victorian and Australian populations. The cultural uniformity within the Shire will be impacted substantially by the projected increase in population of culturally and linguistically diverse migrants and residents of neighbouring localities (Cardinia Shire Council, 2012).

The research involved five local primary schools participating in a racism-reduction initiative (Building Harmony) and an additional local primary school not actively participating in the strategy. Building Harmony is one of few Australian prevention projects to respond proactively to potential diversity issues as new populations arrive (Ozdemir, Soydas, & Grigg, 2012). In

2009 and 2010, the basis of Building Harmony was to inform and develop an appropriate proactive response to potential racial tensions and to establish relationships with the new community and partnerships between the participating schools. A formal pilot test of the project was implemented from November 2010 to November 2011, with positive results including enhancement

Table 1

Mode of Delivery for Building Harmony Activities

Activity	Participants	Time of Delivery	Mode of Delivery
Theatre Games Day	All students	Class time	Five mixed groups
		March 2012	On-campus
Student Leadership Program	Four students from each school	Class time	One mixed group
		May 2012	Off-campus
Shared Literature Activity	All students	Class time	Single-school group
		June 2012	One mixed group On-Campus
Say "No" to Bullying Day	All students	Class time	Five mixed groups
		Parents	One mixed group
Visual Arts Activity	All students	After hours	One mixed group
		July 2012	On-campus
Cross-Cultural Training	Three teachers from each school	Class time	Five mixed groups
		August 2012	On-campus
Cross-Cultural Training	Three teachers from each school	Class time	One mixed group
		September 2012	Off-campus

of school cultural competence and student attitudes towards diversity (Agius, Cooper, Joyce, & Russo, 2011). The activities of the project were based on the available research literature (Cotton, 1993; Greco et al., 2010; Pedersen, Walker, Paradies, & Guerin, 2011; Sanson et al., 1998), but were developed and refined in consultation with the participating schools, in light of the social and emotional developmental level of participants and the community context. The activities included a student leadership program, with students selected as young leaders and diversity champions trained to play an active part in the delivery of the other activities, and taught basic conflict resolution skills to respond to inter-racial conflict. Standardised cross-cultural training was provided to teaching staff in the participating schools. Other activities included a theatre games day, a shared literature activity, a visual arts activity with students working together across cultural and community backgrounds to co-create an artwork, and a 'Say "No" to Bullying' day (see Table 1). Facilitators of diverse backgrounds (including secular, Christian, Catholic, Muslim, Australian background, and non-Australian backgrounds) were involved in delivering learning, as is considered an important aspect of anti-racist education that can assist to break down existing stereotypes.

Participants

The five primary schools involved in the Building Harmony project activities were treated as the intervention group (henceforth known as the Building Harmony Group). One primary school was a government funded secular school, two were non-denominational Christian, one Islamic and one Catholic. Although the intervention schools were involved in the pilot test of Building Harmony in 2010 to 2011, students participating in the 2012 evaluation had not been involved with the project previously. An additional government funded primary school not previously or currently involved in Building Harmony was recruited as the Control Group school. In total, 296 students (9-13 years; 151 Males and 120 Females)

enrolled in years five or six participated in the research.

Procedure

Ethics approval was received by Monash University Human Research Ethics Committee. The authors became involved with Building Harmony participating schools when the project Executive Committee approached them to evaluate the 2012 program activities. The principal of the Control Group school was approached by the authors directly to request participation in the assessment for evaluation purposes. All participating schools obtained permission for students to participate in the research from parents, with no parent declining their child's participation.

The first Building Harmony activities began soon after the beginning of the school year (March 2012), prior to the commencement of the study; pre-test data were collected in April 2012. Post-testing took place in September 2012, two weeks after the final activity, and test-retesting with the project schools took place two weeks later, as recommended in the literature (Pedhazur & Schmelkin, 1991). Although attenuation of an intervention effect may occur swiftly post-intervention, a test-retest assessment was utilised to provide additional validity evidence for RACES. All questionnaires were completed under the supervision of school teachers during class time. Students in the project schools completed the survey online (20-30 minutes to complete); students in the Control Group completed the survey in hard copy (45-60 minutes to complete). The discrepancy in completion time by method of administration is possibly an artefact of young people's increasing reliance on computers and less so on hand writing, thus increasing the time taken to complete the hard copy surveys. At each assessment period, all surveys were completed within 10 days of the first participant completing their survey.

Measures

A demographic questionnaire was utilised to gather general demographic information.

RACES. RACES was completed by respondents in its preliminary 40-item form, but all analyses are based on the final 24-item scale (Grigg & Manderson, 2014a). RACES consists of three interdependent subscales: Racist Attitudes Scale (RAS), an 8-item scale of attitudes reflecting out-group derogation; Accepting Attitudes Scale (AAS), a 12-item scale of attitudes reflecting out-group endorsement; and Ethnocentric Attitudes Scale (EAS), a 4-item scale of attitudes reflecting in-group favouritism (Grigg & Manderson, 2014a). RACES was developed based on qualitative research with Australians of diverse backgrounds and a review of the academic literature, enabling conceptualisations and experiences of real people to complement academic understandings of racism in Australia. A robust and consistent factor structure underlies RACES, demonstrated using both Classical Testing Theory and Item Response Theory. The subscales are appropriately interrelated, and the relationships with RACES and psychopathic personality traits, an existing measure of racism in Australia and a measure of social desirability, have been established. Items in RACES are responded to on a five-point Likert-type scale ranging from “Strongly Disagree” to “Strongly Agree”. Half are reverse scored so higher scores indicate higher levels of acceptance or lower levels of racist attitudes. RACES is internally consistent (total scale and subscale Alpha Coefficient’s range from .79-.91) and possesses factorial and construct validity in children, adolescents and adults, as well as convergent and discriminant validity in adolescents and adults (Grigg & Manderson, 2014a, 2014b, 2014c).

MCSDS-A. A 10-item version of the Marlowe-Crowne Social Desirability Scale (MCSDS; Strahan & Gerbasi, 1972), previously amended for use with Australian youth (MCSDS-A; anonymised for review), was completed. MCSDS-A was used to assess for socially desirable responding, which is integral when measuring sensitive, potentially uncomfortable, or anxiety provoking topics, such as those related to

racism (Anastasi & Urbina, 1996; Janus, 2010; Loewenthal, 2001; Phillips & Ziller, 1997). The 10-item MCSDS is the best shortened version and improves prior forms, including the original 33-item form (Fischer & Fick, 1993). The MCSDS-A is internally consistent (Alpha: .73) and possesses adequate convergent validity, as demonstrated by established relationships with two measures of racist attitudes and a measure of psychopathic personality traits (anonymised for review). The MCSDS-A is responded to on a five-point Likert-type scale ranging from “Strongly Disagree” to “Strongly Agree”. Again, half of the items are reverse scored so higher scores indicate higher levels of social desirability.

SDQ. The 25 core items of the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) were included to evaluate the indirect effects of Building Harmony and to provide evidence of convergent and discriminant validity for RACES. This instrument, a behavioural screening questionnaire designed for use with 3-16 year olds, assesses emotional symptoms (ESS), conduct problems (CPS), hyperactivity/attention symptoms (HAS), peer relationship problems (PPS), and prosocial behaviour (PSS). Each area forms a five-item subscale and the four problematic construct subscales sum to a total difficulties score (TDS), with item response on a three point Likert-type scale ranging from “Not True” to “Certainly True”; 10 are reverse scored so higher scores indicate greater difficulties.

Analytical Techniques

Identifying the factors responsible for change and the nature of the change (i.e., summative or multiplicative) in large-scale interventions can be challenging (Mullan et al., 2014). Research adopting both correlational and experimental designs can assist in determining the behaviour change mechanisms at work (Mullan et al., 2014). In the present research, we examined both the longitudinal impact of the intervention and explored the cross-sectional relationships between the variables at each assessment period.

Table 2

Descriptive Statistics for Building Harmony Project Sample

		Total	Actual	Response Rate	
Pre-Test Sample Size	Building Harmony	116	114	98%	
	Control	180	129	72%	
	Total	296	243	82%	
Post-Test Sample Size	Building Harmony	116	92	79%	
	Control	180	121	67%	
	Total	296	213	72%	
Test-Retest Sample Size	Building Harmony	116	68	59%	
Overall Sample Size	Building Harmony	116	115	99%	
	Control	180	157	87%	
	Total	296	272	92%	
		<i>M</i>	<i>SD</i>	Range	<i>N</i>
Pre-Test	Building Harmony	11.14	0.58	10-12	114
	Control	10.87	0.72	9-13	157
	Total	10.98	0.67	9-13	271
Age (Years)	Building Harmony	11.45	0.64	10-13	114
	Control	11.26	0.74	10-13	157
	Total	11.34	0.71	10-13	271
Test-Retest	Building Harmony	11.49	0.63	10-13	114
		Male	Female	<i>N</i>	
Gender	Building Harmony	74 (65%)	40 (35%)	114	
	Control	77 (49%)	80 (51%)	157	
	Total	151 (56%)	120 (44%)	272	

Table 2 continued:

		Australia	Other	<i>N</i>
Country of Birth	Building Harmony	95 (83%)	20 (17%)	115
	Control	142 (90%)	15 (10%)	157
	Total	237 (87%)	35 (13%)	272
		Australia	Other	<i>N</i>
Building Harmony	Mother	67 (56%)	48 (42%)	115
	Father	49 (43%)	65 (57%)	114
Control	Mother	112 (74%)	39 (26%)	151
	Father	115 (77%)	35 (23%)	150
Total	Mother	179 (67%)	87 (33%)	266
	Father	164 (62%)	100 (38%)	264

Note. Various participants did not provide date, country, or parent country of birth data.

Results

Pre-test, post-test, and test-retest data were cleaned and analysed using SPSS 20.0. Several datasets were created to enable differential data treatment for (1) comparison of Control and Building Harmony groups on the examined variables (i.e., evaluation of the effect of Building Harmony), (2) evaluation of the strength of relationships between examined variables, and (3) evaluation of the test-retest reliability of RACES. Below, we outline preliminary data analyses, present the results for RACES validation and then present the results of the Building Harmony initiative. Descriptive statistics are shown in Table 2.

Preliminary Data Analysis

Evaluation of RACES psychometric properties. All data across each assessment period and both groups was combined. A missing data analysis was performed and all

cases with 5% or more data missing across RACES, MCSDS-A, or SDQ were removed. Separate regression analyses for each of the alternate scales (RACES, MCSDS-A, and SDQ) were used to deal with remaining missing data. Data were then recombined into a single data set to maximise the sample size for analysis.

The test-retest assessment data provided by the Building Harmony Group was analysed separately, with post-test and test-retest data collated and a missing data analysis performed. All cases with 5% or more data missing across RACES, MCSDS-A, or SDQ were removed in addition to all cases missing either post-test or test-retest data. For the test-retest correlation analyses, only the test-retest data was utilised. Again, all cases with 5% or more data missing independently across RACES, MCSDS-A, or SDQ were removed. There were no remaining cases with less than 5% missing

data, but some cases had full data for RACES and MCSDS-A and no data for the SDQ. Data were recombined into a single data set to maximise the sample size for analysis. For all test-retest assessment data analyses, casewise deletion was used to deal with any unpaired data. All analyses in the evaluation of the RACES were conducted utilising the General Linear Model (GLM) approach to data analysis.

Building Harmony evaluation. Missing data analysis was performed and all cases with 5% or more data missing across RACES, MCSDS-A, or SDQ, and all cases with only pre- or post-test data, were removed. Intention to treat analysis to carry forward data was inappropriate due to the lack of connection between the intervention and the absence of students at post-test (i.e., students were absent due to illness). Linear Mixed Models (LMM) were used to deal with all remaining missing data because of its greater flexibility than the traditional GLM and its capacity for a more nuanced analysis of intervention data (Schneider, 2013). Due to missing data, only 41 respondents from the Building Harmony Group (75 responses omitted due to missing data at pre- and/or post-test) and 84 respondents from the Control Group (96 responses omitted due to missing data at pre- and/or post-test) were included in analyses at both pre- and post-test. LMM was used to assess the effect of the intervention on RACES, MCSDS-A and SDQ scores.

For all analyses, assumptions were examined via the inspection of normality plots. Although some variables appeared non-normal (i.e., mild to moderate skew), the sample size was large enough for the selected statistical analyses to be robust; transforming non-normal data is a questionable practice and for sample sizes above 30, the sampling distribution of the mean can be safely assumed to be normal (Field, 2009; Games, 1984; Salkind, 2006). No significant univariate or multivariate outliers were detected.

Evaluation of RACES Convergent and Discriminant Validity

The relationship between racial, ethnic, cultural, and religious acceptance, and the various strengths and difficulties experienced by the students, was assessed across both groups and all data collection periods. Due to consistency across assessment periods, only the pre-test correlation analysis is shown (see Table 3). No significant correlations had 95% CIs that crossed zero for either group, the total sample, or across pre- and post-test. The 95% CIs effectively spanned one strength rating at their widest point (i.e., from trivial to small; small to moderate; moderate to large, etc.) (Cohen, 1988). As most significant correlations were moderate to large, these findings provide confidence that many of the relationships were meaningful. *RACES overall scale correlation findings.*

For both groups and all assessment periods, there was a significant positive relationship with each of the subscales, with the MCSDS-A, and with the PSS. Conversely, across groups and assessment periods, there was a significant negative relationship with the TDS, and the HAS and CPS. There was an inconsistent relationship between the overall RACES and the PPS and no significant relationship between the overall RACES and the ESS across either group or assessment period.

The consistency of the RACES relationship with its subscales provides important evidence for the overall construct validity of the measure. However, the consistency of the relationship with the MCSDS-A, suggests that overall RACES scores are related to socially desirable responding. The RACES scores therefore may not represent the true attitudes of the participants. However, previous scale developers have suggested a correlation in the range of $> .60$ to be indicative of excessive response bias contamination, a level not reached in the current research (Lee, Gibbons, Thompson, & Timani, 2009). The consistency of the RACES relationship with the PSS, TDS, HAS, and CPS provides important evidence for the convergent

Table 3

Correlation Analyses for Pre-Test Data

	2.		3.		4.		5.		6.		7.		8.		9.		10.		11.			
	BH	CG	TS	BH	CG	TS	BH	CG	TS	BH	CG	TS	BH	CG	TS	BH	CG	TS	BH	CG	TS	
1.	<i>r</i>	.93***	.85***	.88***	.80***	.81***	.76***	.76***	.76***	.76***	.76***	.76***	.76***	.76***	.76***	.76***	.76***	.76***	.76***	.76***	.76***	.76***
RACES	<i>p</i>	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
2.	<i>r</i>	.57***	.46***	.48***	.61***	.50***	.54***	.59***	.28*	.42***	.28*	.28*	.28*	.28*	.28*	.28*	.28*	.28*	.28*	.28*	.28*	.28*
2.AAS	<i>p</i>	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
3.	<i>r</i>	.49***	.53***	.52***	.44***	.11	.17*	.21	.27***	.30***	-.07	-.02	-.08	-.38**	-.14	-.25***	-.25*	-.06	-.10	-.32**	-.18*	-.27***
3.RAS	<i>p</i>	<.001	<.001	<.001	<.001	.23	.015	.055	.003	<.001	.52	.85	.26	.001	.13	<.001	.024	.53	.14	.004	.049	<.001
4.	<i>r</i>	.39***	.23*	.28***	-.11	-.28**	-.23**	-.26*	-.19*	-.22**	-.21	-.36***	-.32***	-.26*	-.30**	-.27***	-.30**	-.39***	-.36***	-.34**	-.26**	-.29***
4.EAS	<i>p</i>	<.001	.010	<.001	.32	.002	.001	.021	.042	.002	.058	<.001	<.001	.020	.001	<.001	.007	<.001	<.001	<.001	.002	.005
5.	<i>r</i>	-.13	-.30**	-.20**	.05	-.02	-.02	-.22*	-.39***	-.29***	-.30**	-.25**	-.28***	-.20	-.33***	-.25***	-.41***	.28***	-.20	-.33***	-.25***	-.41***
5.PSS	<i>p</i>	.24	.001	.006	.66	.81	.74	.047	<.001	<.001	.006	.005	<.001	.069	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
6.	<i>r</i>	.44***	.23**	.33***	.37**	.44***	.43***	.25*	.17	.19***	.77***	.70***	.74***	.43***	.42***	.41***	.41***	.26*	.34***	.32***	.27*	.35***
6.HAS	<i>p</i>	<.001	.010	<.001	.001	<.001	.027	.057	.007	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
7.	<i>r</i>	.26*	.34***	.32***	.27*	.35***	.30***	.74***	.72***	.75***	-.09	-.06	-.07	.019	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
7.ESS	<i>p</i>	.019	<.001	<.001	.015	<.001	<.001	.015	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
8.	<i>r</i>	.26*	.39***	.32***	.66***	.75***	.72***	.56***	.39***	.46***	.021	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
8.CPS	<i>p</i>	.021	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
9.	<i>r</i>	.61***	.60***	.58***	-.09	-.06	-.08	-.09	-.06	-.08	-.09	-.06	-.08	-.09	-.06	-.08	-.09	-.06	-.08	-.09	-.06	-.08
9.PPS	<i>p</i>	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
10.	<i>r</i>	.41***	.34***	.37***	-.41***	.34***	.37***	-.41***	.34***	.37***	-.41***	.34***	.37***	-.41***	.34***	.37***	-.41***	.34***	.37***	-.41***	.34***	.37***
10.TDS	<i>p</i>	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
11.	<i>r</i>																					
11.MGSDS-A	<i>p</i>																					

Note. RACES = Racism, Acceptance, and Cultural-Ethnocentrism Scale; AAS = Accepting Attitudes Scale; RAS = Racist Attitudes Scale; EAS = Ethnocentric Attitudes Scale; PSS = SDQ Prosocial Scale; HAS = SDQ Hyperactivity Scale; ESS = SDQ Emotional Symptoms Scale; CPS = SDQ Conduct Problems Scale; PPS = SDQ Peer Problems Scale; TDS = SDQ Total Difficulties score; MGSDS-A = Marlowe Crowne Social Desirability Scale Australian; BH = Building Harmony Group; CG = Control Group; TS = Total Sample.
 * *p* < .05. ** *p* < .01. *** *p* < .001.

validity of the scale. As expected, higher levels of acceptance are related to higher levels of prosocial behaviour and lower levels of social, behavioural, and emotional difficulties.

The inconsistent and non-significant relationship respectively between the overall RACES and the PPS and ESS contradicts the expectation that difficulties with peers and expression of emotion would be related to lower levels of acceptance of diversity. As detailed below, these findings reinforce prior suggestions that the RACES subscales may be more meaningful when utilised interdependently (i.e., as a three factor scale), rather than as three independent scales (Grigg & Manderson, 2014a, 2014c).

Accepting Attitudes Scale correlation findings. For the AAS across both groups and assessment periods, there was a significant positive relationship with the RAS, with the MCSDS-A, and with the PSS. Conversely, across both groups and assessment periods, there was a significant negative relationship with the HAS. There was no significant relationship with the ESS across either group or assessment period, and an inconsistent relationship between the AAS and the EAS, CPS, PPS, and TDS.

Racist Attitudes Scale correlation findings. Across both groups and assessment periods, there was a significant positive relationship with the EAS and with the MCSDS-A, and a significant negative relationship with the TDS. There was no significant relationship with the ESS across either group or assessment period, and an inconsistent relationship between the RAS and the PSS, HAS, CPS, and PPS.

Ethnocentric Attitudes Scale correlation findings. There was no variable with a significant positive or negative relationship across both groups and assessment periods. There was no significant relationship with the ESS across either group or assessment period. There was a consistent negative relationship with the HAS and CPS for both the Control Group and the Total Sample, but these relationships were absent for the Building Harmony Group across assessment periods. There was an

inconsistent relationship between the EAS and the PSS, PPS, TDS, and MCSDS-A.

MCSDS-A correlation findings. The MCSDS-A findings were the most stable and predictable of all the measured variables. Across both groups and assessment periods, there was a significant positive relationship with the RACES total scale, AAS, RAS, and PSS, and a significant negative relationship with the HAS, CPS, and TDS. There was no significant relationship with the ESS or CPS for either group across assessment periods. The EAS was inconsistent for the Building Harmony Group, but was significantly positively associated with the Control Group and the Total Sample across assessment periods.

SDQ correlation findings. The SDQ total scale and subscale correlation findings may further explain some of the inconsistent findings reported above. Only the CPS demonstrated significant relationships in the expected direction with the TDS or other SDQ subscales across both groups and assessment periods. The inconsistent findings with the SDQ total scale and subscales may have caused some of the discrepancies for the RACES total scale and subscale relationships.

RACES convergent and discriminant validity findings summary. The positive relationships between desirable variables and negative relationships between undesirable variables were expected. The inconsistent relationship between the EAS and MCSDS-A may be partly due to the limited length of the EAS (i.e., four items). The lack of a significant relationship with the ESS and CPS may partly explain the inconsistent findings reported above. Overall, the correlation findings support the construct, convergent, and discriminant validity of the RACES total scale, RACES subscales, and the MCSDS-A, with most relationships consistent and in the expected direction.

Evaluation of RACES Test-Retest Reliability

Eleven paired-samples t-tests and Pearson correlations were conducted to assess the stability of RACES, SDQ, and MCSDS-A total scales and subscales (tabular results available upon request). The t

-test results indicate that each of the RACES, SDQ, and MCSDS-A total scales and subscales were of acceptable stability indicated by the lack of a statistically significant difference between post- and test-retest. The correlation results suggest that each of the RACES, SDQ, and MCSDS-A total scales were of acceptable stability (correlations above .70) (Nunnally, 1978). However, the EAS was of less than desirable stability, reinforcing the importance of utilising the three RACES subscales interdependently rather than independently. The CPS and PPS were also of poor stability,

which may have contributed to the inconsistent results reported above.

Building Harmony Evaluation

A series of 2x2 ANOVAs with Group (Building Harmony Group and Control Group) and Assessment (Pre-Test and Post-Test) as the two factors were conducted to assess the effect of the intervention on RACES, SDQ, and MCSDS-A total scale and subscale scores (tabular results available upon request). Due to the significant increase in MCSDS-A scores for the Control Group from pre- to post-test, a series of 2x2 ANCOVAs were conducted with the same

Table 4: Descriptive Statistics for RACES and MCSDS-A with and without Covariance

		RACES			RACES (MCSDS-A Controlled)			MCSDS-A			
		Building Harmony	Control	Total	Building Harmony	Control	Total	Building Harmony	Control	Total	
Pre- Test	<i>M</i>	74.83	67.81	71.32	74.83	67.88	71.35	28.76	27.66	28.21	
	<i>SE</i>	1.59	1.13	0.98	1.59	1.16	0.98	0.66	0.47	0.40	
	<i>df</i>	169.66	173.91	171.09	161.45	172.57	165.33	182.29	189.32	184.70	
Post - Test	<i>M</i>	75.51	71.64	73.58	75.51	72.25	73.88	27.73	28.96	28.35	
	<i>SE</i>	1.59	1.18	0.99	1.50	1.22	1.00	0.66	0.48	0.41	
	<i>df</i>	169.66	188.78	176.52	161.45	187.15	171.19	182.29	192.15	185.72	
		Accepting Attitudes Scale			Racist Attitudes Scale			Ethnocentric Attitudes Scale			
		Building Harmony	Control	Total	Building Harmony	Control	Total	Building Harmony	Control	Total	
No CV	Pre- Test	<i>M</i>	39.61	37.79	38.70	22.27	18.13	20.20	12.95	11.78	12.36
		<i>SE</i>	0.87	0.61	0.53	0.70	0.50	0.43	0.39	0.28	0.24
		<i>df</i>	183.40	184.82	183.87	194.81	197.56	195.72	183.19	184.67	183.68
CV	Post- Test	<i>M</i>	40.05	38.66	39.36	22.12	19.66	20.89	13.34	12.85	13.10
		<i>SE</i>	0.87	0.62	0.53	0.70	0.50	0.43	0.39	0.28	0.24
		<i>df</i>	183.40	193.50	186.88	194.81	200.30	196.67	183.19	187.65	184.68
CV	Pre- Test	<i>M</i>	39.61	37.90	38.75	22.27	18.15	20.21	12.95	11.77	12.36
		<i>SE</i>	0.87	0.62	0.53	0.70	0.51	0.44	0.39	0.28	0.24
		<i>df</i>	173.14	181.44	175.99	187.25	196.24	190.42	174.74	183.40	177.71
CV	Post- Test	<i>M</i>	40.05	39.16	39.61	22.12	19.70	20.91	13.34	12.96	13.15
		<i>SE</i>	0.87	0.64	0.54	0.70	0.52	0.44	0.94	0.29	0.24
		<i>df</i>	173.14	191.36	179.76	187.25	201.23	192.40	174.74	187.82	179.33

Note. CV = inclusion of covariate. No CV = no inclusion of covariate.

Table 5: Descriptive Statistics for Individual SDQ Subscales with and without Covariance

			Prosocial Scale			Hyperactivity Scale			Emotional Symptoms Scale		
			BH	CG	TS	BH	CG	TS	BH	CG	TS
No	Pre-Test	<i>M</i>	8.07	8.31	8.19	3.51	4.57	4.04	2.76	3.26	3.01
		<i>SE</i>	0.27	0.19	0.16	0.39	0.27	0.24	0.39	0.27	0.24
		<i>df</i>	182.01	183.47	182.49	177.85	180.78	178.82	175.28	175.28	175.28
CV	Post-Test	<i>M</i>	7.95	8.49	8.22	3.44	4.45	3.95	3.17	3.32	3.25
		<i>SE</i>	0.27	0.19	0.16	0.39	0.27	0.24	0.39	0.27	0.24
		<i>df</i>	182.01	182.01	182.01	177.85	177.85	177.85	175.28	175.28	175.28
CV	Pre-Test	<i>M</i>	8.07	8.29	8.18	3.51	4.63	4.07	2.76	3.29	3.03
		<i>SE</i>	0.27	0.19	0.17	0.39	0.28	0.24	0.39	0.28	0.24
		<i>df</i>	178.74	187.30	181.69	175.42	185.68	178.97	163.00	169.93	165.33
CV	Post-Test	<i>M</i>	7.95	8.44	8.20	3.44	4.52	3.98	3.17	3.32	3.25
		<i>SE</i>	0.27	0.19	0.17	0.39	0.28	0.24	0.39	0.28	0.24
		<i>df</i>	178.74	188.74	182.21	175.42	185.68	178.97	163.00	172.83	166.34
			Conduct Problems Scale			Peer Problems Scale			Total Difficulties		
			BH	CG	TS	BH	CG	TS	BH	CG	TS
No	Pre-Test	<i>M</i>	1.59	2.35	1.97	1.78	1.66	1.72	9.63	11.79	10.71
		<i>SE</i>	0.28	0.20	0.17	0.25	0.17	0.15	0.93	0.65	0.57
		<i>df</i>	195.58	195.58	195.58	216.11	217.16	216.46	184.42	187.39	185.40
CV	Post-Test	<i>M</i>	2.02	2.05	2.04	2.17	1.61	1.89	10.81	11.52	11.16
		<i>SE</i>	0.28	0.20	0.17	0.25	0.17	0.15	0.93	0.65	0.57
		<i>df</i>	195.58	197.00	196.05	216.11	216.11	216.11	184.42	185.90	184.91
CV	Pre-Test	<i>M</i>	1.59	2.39	1.99	1.78	1.65	1.71	9.63	11.89	10.76
		<i>SE</i>	0.28	0.20	0.17	0.25	0.18	0.15	0.91	0.66	0.56
		<i>df</i>	193.03	199.41	195.24	209.20	214.31	211.02	182.68	192.48	186.11
CV	Post-Test	<i>M</i>	2.02	2.13	2.08	2.17	1.60	1.89	10.81	11.66	11.23
		<i>SE</i>	0.28	0.21	0.17	0.25	0.18	0.15	0.91	0.67	0.56
		<i>df</i>	193.03	203.15	196.65	209.20	215.12	211.34	182.68	182.68	186.63

Note. BH = Building Harmony Group. CG = Control Group. TS = Total Sample.

factors as above and with MCSDS-A as a covariate to assess the effect of the intervention on RACES and SDQ total scale and subscale scores to control for the effect of socially desirable responding. Descriptive statistics for the overall scales and subscales (controlling for the MCSDS-A and without

MCSDS-A as a covariate) and the results of the analyses are set out in Tables 4-7.

Significant Pre- to Post-Test effects indicate that a group had a significant change from pre- to post-test. Significant Group effects indicate that the groups were significantly different when averaged across the two assessment periods. Significant Assessment effects indicate that scores were

Table 6
Pre-Test to Post-Test ANOVA and ANCOVA Results for RACES and SDQ Total Scale and Subscales and MCSDS-A by Group

	Accepting Attitudes Scale			Racist Attitudes Scale			Ethnocentric Attitudes Scale			RACES Total Score			MCSDS-A											
	ANOVA	ANCOVA	CG	ANOVA	ANCOVA	CG	ANOVA	ANCOVA	CG	ANOVA	ANCOVA	CG	ANOVA	ANCOVA	CG									
F	0.29	2.20	0.32	4.35	0.04	8.96	0.04	7.81	1.14	16.97	1.16	18.82	0.24	12.83	0.25	15.33	2.68	7.74	-	-				
d	0.08	0.15	0.08	0.22	0.03	0.33	0.03	0.33	0.16	0.42	0.08	0.46	0.07	0.36	0.07	0.40	0.24	0.30	-	-				
9 L	-0.35	-0.15	-0.35	-0.09	-0.40	0.03	-0.40	0.02	-0.28	0.11	-0.35	0.15	-0.37	0.06	-0.36	0.09	-0.19	-0.01	-	-				
5 L																								
%																								
C U	.51	0.45	0.51	0.52	0.47	0.64	0.47	0.63	0.59	0.72	0.52	0.76	0.50	0.66	0.50	0.70	0.68	0.60	-	-				
I L																								
d	116.26	120.31	105.94	115.92	119.18	103.59	115.00	117.24	119.27	106.71	114.68	106.10	113.74	95.73	108.89	111.55	118.17	-	-	-				
f																								
p	.59	.14	.58	.39	.84	.003**	.84	.006**	.29	<.001***	.28	<.001***	.63	.001**	.62	<.001***	.11	.006**	-	-				
	Prosocial Scale			Hyperactivity Scale			Emotional Symptoms Scale			Conduct Problems Scale			Peer Problems Scale			SDQ Total Difficulties								
	ANOVA	ANCOVA	CG	ANOVA	ANCOVA	CG	ANOVA	ANCOVA	CG	ANOVA	ANCOVA	CG	ANOVA	ANCOVA	CG	ANOVA	ANCOVA	CG	ANOVA	ANCOVA	CG			
F	0.26	1.18	0.24	0.66	0.05	0.25	0.05	1.58	0.07	1.78	0.01	2.41	2.28	2.23	1.35	1.96	0.07	1.90	0.05	1.80	0.19	1.70	0.11	
d	0.07	0.10	0.07	0.09	0.03	0.05	0.03	0.04	0.16	0.02	0.16	0.01	0.24	0.16	0.24	0.14	0.24	0.03	0.24	0.03	0.20	0.05	0.20	
9 L	-0.36	-0.20	-0.36	-0.22	-0.41	-0.25	-0.41	-0.26	-0.27	-0.28	-0.29	-0.20	-0.14	-0.19	-0.17	-0.19	-0.27	-0.19	-0.27	-0.19	-0.27	-0.63	-0.26	-0.23
5 L																								
%																								
C U	0.50	0.41	0.50	0.39	0.46	0.35	0.46	0.35	0.60	0.33	0.60	0.31	0.67	0.47	0.67	0.44	0.68	0.33	0.68	0.33	0.24	0.35	0.63	
I L																								
d	122.34	122.81	110.08	117.10	120.67	121.58	120.67	121.58	123.00	115.17	120.87	121.42	107.91	116.08	122.47	123.10	113.96	123.04	118.10	119.63	105.67	114.49	-	-
f																								
p	.61	.28	.42	.63	.83	.62	.83	.62	.21	.80	.19	.91	.12	.13	.25	.16	.79	.17	.82	.18	.67	.20	.74	

Note. BH = Building Harmony Group. CG = Control Group. Numerator $df = 1$ for all analyses.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 7

Overall ANCOVA Results for RACES and SDQ Total Scale and Subscales

	Accepting Attitudes Scale				Racist Attitudes Scale				Ethnocentric Attitudes Scale				RACES Total Score				SDQ Total Difficulties								
	G	A	GA	Pr	G	A	GA	Pr	G	A	GA	Pr	G	A	GA	Pr	G	A	GA	Pr					
<i>F</i>	1.87	2.97	0.70	2.58	0.68	19.32	2.37	3.47	22.46	7.66	3.33	12.13	3.11	5.96	0.62	8.25	8.27	4.41	12.50	2.66	2.52	0.70	1.55	3.99	0.57
<i>d</i>	0.26	0.14	0.30	0.15	0.83	0.14	0.89	0.52	0.35	0.29	0.46	0.09	0.54	0.23	0.66	0.30	0.30	0.07	0.38	0.14	0.38	0.14	0.00	-0.23	
9 L	-0.12	-0.11	-0.07	-0.22	0.43	-0.11	0.50	0.14	-0.03	0.04	0.08	-0.28	0.16	-0.02	0.28	-0.07	-0.08	-0.17	0.00	-0.23					
5 L																									
%U	0.63	0.39	0.68	0.53	1.21	0.39	1.28	0.89	0.72	0.54	0.84	0.47	0.91	0.48	1.04	0.68	0.67	0.32	0.75	0.51					
C L																									
I																									
<i>d</i>	120.44	109.60	109.60	175.99	179.76	115.75	107.67	107.67	190.42	192.40	117.93	109.55	109.55	177.71	179.33	116.49	100.83	100.83	165.33	171.19	116.21	108.79	108.79	186.63	
<i>f</i>	.17	.088	.41	.11	<.001***	.13	.065	<.001***	.006**	.071	.001**	.081	.016*	.43	.005**	.005**	.038*	.001**	.11	.12	.41	.22	.047*	.45	

	Prosocial Scale				Hyperactivity Scale				Emotional Symptoms Scale				Conduct Problems Scale				Peer Problems Scale								
	G	A	GA	Pr	G	A	GA	Pr	G	A	GA	Pr	G	A	GA	Pr	G	A	GA	Pr					
<i>F</i>	1.46	0.01	0.77	0.42	2.18	6.79	0.17	0.01	5.00	4.58	0.63	1.29	1.00	1.29	0.10	2.39	0.25	3.59	5.38	0.10	2.00	0.94	1.54	0.19	3.49
<i>d</i>	0.23	0.01	0.13	0.28	0.49	0.03	0.44	0.42	0.15	0.08	0.21	0.06	0.29	0.05	0.44	0.06	0.26	0.10	0.08	0.35	0.08	0.35	0.08	0.35	
9 L	-0.15	-0.24	-0.25	-0.09	0.11	-0.21	0.06	0.04	-0.22	-0.17	-0.17	-0.31	-0.09	-0.20	0.06	-0.32	-0.12	-0.15	-0.29	-0.03					
5 L																									
%U	0.60	0.26	0.50	0.66	0.87	0.28	0.81	0.80	0.52	0.33	0.58	0.43	0.66	0.30	0.81	0.43	0.63	0.35	0.45	0.72					
C L																									
I																									
<i>d</i>	119.66	112.56	112.56	181.69	182.21	116.94	109.22	109.22	178.82	177.85	119.78	111.86	111.86	165.33	166.34	116.44	110.77	110.77	195.24	196.65	121.40	117.11	117.11	211.34	
<i>f</i>	.23	.92	.38	.52	.14	.010*	.68	.93	.027*	.034*	.43	.26	.32	.26	.75	.13	.62	.061	.021*	.76	.16	.33	.22	.66	

Note. G = Group Effect, A = Assessment Effect, GA = Group by Assessment Effect, Pr = Pre-Test Effect, Po = Post-Test Effect, Numerator *df* = 1 for all analyses.

* $p < .05$, ** $p < .01$, *** $p < .001$.

significantly different over pre- and post-test when averaged across the two groups. Significant Group by Assessment effects indicate that the two groups had a significantly different change in scores from pre-test to post-test. Significant Pre/Post-Test effects indicate that the groups were significantly different at pre/post-test. Confidence intervals in the analyses were relatively wide, revealing the imprecision of the analyses and highlighting the need for replication with a larger sample size and decreased variability before firm conclusions can be drawn (Cumming, Fidler, Kalinowski, & Lai, 2012). The most reliable effects were able to be examined by focusing only on the effects with 95% CIs that did not drop below zero.

Of the significant Pre- to Post Test effects, the Control Group MCSDS-A had a lower bound 95% CI below zero and so can be discounted. However, the Control Group had a significant increase in RACES Total Score and both the RAS and EAS subscales both with and without MCSDS-A as a covariate. No other significant Pre- to Post-Test effects were found.

Of the significant Pre-Test effects, the HAS and CPS (without MCSDS-A as a covariate) had lower bound 95% CIs below zero and so they can be discounted. In contrast, RAS, EAS, and RACES total score were significantly higher for the Building Harmony Group at pre-test both with and without MCSDS-A as a covariate. The HAS, CPS, and TDS were each significantly higher for the Control Group with the MCSDS-A as a covariate only. No other significant Pre-Test effects were found.

The RAS and HAS were significantly higher for the Building Harmony Group and Control Group respectively at post-test both with and without MCSDS-A as a covariate. No other significant Post-Test effects were found and none of the significant Post-Test effects had a lower bound 95% CI below zero.

The RAS, RACES total score, and HAS were found to have a significant Group effects both with and without MCSDS-A as a covariate. The EAS without MCSDS-A as a

covariate was found to have a non-significant Group effect, but this effect had a lower bound 95% CI above zero. No other significant Group effects were found and none of the significant Group effects had a lower bound 95% CI below zero.

Of the significant Assessment effects, the RACES total score (with and without MCSDS-A as a covariate) had lower bound 95% CIs below zero and so can be discounted. In contrast, the EAS was found to have a significant Assessment effect both with and without MCSDS-A as a covariate. No other significant Assessment effects were found.

As Group by Assessment effects are interaction effects, effect size estimate are unavailable. A significant Group by Assessment effect was found for MCSDS-A and CPS without MCSDS-A as a covariate. A significant Group by Assessment effect was also found for RACES total score with MCSDS-A as a covariate. No other significant Group by Assessment effects were found.

Building Harmony evaluation summary.

Overall, many effects were trivial, however, some effects entered the moderate to large range, with lower bound 95% CIs in the small to moderate range (Cohen, 1988). If focus is targeted only upon the non-trivial effects significant meaning can be drawn from the data. Results provide tenuous efficacy evidence for the Building Harmony initiative in enhancing racial attitudes and social, emotional, and behavioural strengths. *Comparison with Building Harmony Pilot Test*

Due to the tenuous efficacy evidence for the Building Harmony initiative, results from the 2011 Building Harmony pilot test were examined. Twenty-four single-sample t-tests were conducted on de-identified data from the 2011 Building Harmony project to assess the difference between the mean SDQ subscale and total difficulties scores from the 2011 Building Harmony initiative post-test and the current sample. These results demonstrated that the Control Group at pre- and post-test had significantly more social, emotional, and behavioural difficulties than

the 2011 Building Harmony participants at post-test (HAS, EAS, CPS, and TDS). Yet, the Building Harmony Group only had differences for HAS at pre-test and CPS at post-test. Further, the Control Group at post-test had a significantly higher PSS, which again suggests a potential positive change throughout the assessment periods. These findings support the effectiveness of the Building Harmony initiative if presented to a novel audience.

Discussion

The results support the reliability and validity of RACES, with the measure shown to be stable over time and demonstrating expected relationships with social, emotional, and behavioural strengths and difficulties. The overall results of the evaluation of the Building Harmony project were also positive, and support the program in maintaining accepting attitudes toward distinct groups, increasing social, emotional, and behavioural strengths, and reducing social, emotional, and behavioural difficulties. However, the results only equivocally support the efficacy of the initiative, consistent with considerable intergroup relations research conducted in educational institutions (Greco et al., 2010; Walker & Crogan, 1998).

RACES Evaluation

Relationships with RACES scales and subscales were predominantly as expected. Participants with greater social, emotional, and behavioural difficulties were found to have lower levels of acceptance of difference, and the converse also resulted, consistent with previous research. For example, prosocial attitudes including empathic concern are positively related to intention to intervene in a racist situation and empathy is negatively related to racism (Neto & Pedersen, 2013; Pettigrew & Tropp, 2006). A range of social, emotional, and behavioural difficulties, including those measured by the SDQ, have been robustly linked to self-esteem in children and youth (Bolger, Patterson, & Kupersmidt, 1998; Bosacki, Dane, Marini, & YLC-CURA, 2007; Farrell, 1995; Slomkowski, Klein, & Mannuzza, 1995; Swinson, 2008). This is important in

the context of the current research because self-esteem is implicated in intergroup bias. However, the link is inconsistent, with historical research associating lower levels of self-esteem to higher levels of racist attitudes (Bagley, 1979) and more recent reviews suggesting high levels of self-esteem result in intergroup bias and consequently racist attitudes (Aberson, Healy, & Romero, 2000). This general inconsistency and the more recent unintuitive findings may have contributed to the somewhat variable results with RACES and SDQ subscales.

Building Harmony Evaluation

Variable correlational results may have contributed to the tenuous efficacy results for Building Harmony, which had a strong underlying goal of enhancing the self-esteem of participants. Moreover, the Control Group had a significant increase in both RACES total scale, subscales, and MCSDS-A scores from pre-test to post-test, which may indicate socially desirable responding rather than an increase in acceptance of difference. A potential contributor to the increase for the Control Group may also be growth in cultural diversity that occurred throughout the school year, leading to students' greater exposure to and interaction with people from diverse backgrounds (Rutland, Cameron, Bennett, & Ferrell, 2005; Tropp, 2005; cf. Barlow et al., 2012). In contrast, the focus on racism reduction in the Building Harmony intervention activities may have brought about a critical awareness of racism in participants, increasing student knowledge, mindfulness, and consequently self-reporting of their racist behaviours and attitudes (Howarth, 2009).

The pre-test occurred after the implementation of the first Building Harmony intervention activity. The significant difference between the Building Harmony Group and Control Group on RACES scores at pre-test may therefore reflect that the first activity was effective in increasing the Building Harmony Group's acceptance of diversity. Alternatively, there may have been a positive overall effect of participation in Building Harmony on the

school culture: although the study participants were not previously directly involved in the Building Harmony activities, the participating schools had been so engaged for several years. The consistency between the pre- and post-test scores for the Building Harmony Group indicates that the subsequent activities were at least effective in maintaining this level of acceptance. Moreover, despite the apparent increase in RACES total scale and subscale scores for the Control Group, the scores for this group did not reach a level higher than the Building Harmony Group at either assessment period. These interpretations are reinforced by the consistency between RACES and SDQ findings; both scales and their components were (1) relatively stable for the Building Harmony Group and did not significantly increase, and (2) generally more positive for the Building Harmony Group across both assessment periods. Comparison with results from the 2011 Building Harmony initiative further support these conclusions. These findings endorse the effectiveness of the Building Harmony initiative, if presented to a novel audience.

Limitations

Although the results of the research are encouraging, a number of limitations need to be acknowledged. As noted above, due to delays in final ethical approval for the research, the pre-test for the Building Harmony Group occurred after the implementation of the first intervention activity. In addition, although the study participants had not previously been involved in Building Harmony, the participating schools had been engaged with the project for a number of years. These factors suggest that accurate baseline measurement may not have been achieved. Similarly, the final activity (cross-cultural training) involved teachers only. Such training, although integral to school culture, is unlikely to have an immediate impact upon the attitudes of students. The elapsed time between the last activity directly targeting student participants and the completion of the post-test may have limited the overall efficacy results of the program. The substantial number of invalid

responses may also have biased the results. Future replications of Building Harmony might re-order the activities to facilitate the most positive and enduring impact upon participants, whilst ensuring that appropriate pre- and post-testing is completed. Building Harmony was implemented in primary schools from a small community in the southeastern region of Melbourne, Australia, and results may not be generalisable to the wider Australian population. Learnings from Building Harmony could be transferred to, and evaluated in, other settings and initiatives around Australia to develop and refine anti-racism and pro-diversity projects that are efficacious and able to be broadly implemented.

Conclusion

The results provide strong evidence of the internal consistency, test-retest reliability, construct validity, discriminant validity, and convergent validity of RACES as a measure of racial, ethnic, cultural, and religious acceptance, which can be utilised to quantify racist attitudes in intervention evaluations. Moreover, Building Harmony program activities appear effective in maintaining positive attitudes towards diversity, and current levels of social, emotional, and behavioural strengths and difficulties in participants.

An initial purpose of this research was to evaluate the appropriateness, reliability, and validity of RACES to evaluate anti-racism interventions. Strong validity evidence suggests that it is robust and able to be disseminated and utilised in schools implementing similar anti-racism and pro-diversity interventions around Australia. An additional aim was to evaluate the efficacy of the Building Harmony anti-racism and pro-diversity initiative in positively changing the attitudes of participants. Tenuous evidence for the Building Harmony project was found, but further implementation and appraisal of the intervention is warranted. The ongoing evaluation of such programs utilising empirically validated measures, such as RACES, will ensure that more effective racism-reduction initiatives can be

implemented and community levels of racism may be subsequently reduced whilst enhancing the racial, ethnic, cultural, and religious acceptance of Australian society.

References

- Aberson, C. L., Healy, M., & Romero, V. (2000). Ingroup bias and self-esteem: A meta-analysis. *Personality and Social Psychology Review*, 4(2), 157-173. doi: 10.1207/s15327957pspr0402_04
- Aboud, F. E., & Doyle, A.-B. (1996). Parental and peer influences on children's racial attitudes. *International Journal of Intercultural Relations*, 20(3-4), 371-383. doi: [http://dx.doi.org/10.1016/0147-1767\(96\)00024-7](http://dx.doi.org/10.1016/0147-1767(96)00024-7)
- Aboud, F. E., Tredoux, C., Tropp, L. R., Brown, C. S., Niens, U., & Noor, N. M. (2012). Interventions to reduce prejudice and enhance inclusion and respect for ethnic differences in early childhood: A systematic review. *Developmental Review*, 32(4), 307-336. doi: 10.1016/j.dr.2012.05.001
- Agius, C., Cooper, T., Joyce, A., & Russo, A. (2011). *The Building Harmony in the Cardinia Growth Corridor findings report 2010-2011*. Melbourne, VIC: Windermere Child and Family Services.
- Anastasi, A., & Urbina, S. (1996). *Psychological testing (7th ed.)*. Upper Saddle River, NJ: Prentice Hall.
- Australian Bureau of Statistics. (2008). 3222.0 - Population projects, Australia, 2006 to 2101. Retrieved from <http://www.abs.gov.au/Ausstats/abs@.nsf/mf/3222.0>.
- Bagley, C. (1979). Self-esteem as a pivotal concept in race and ethnic relations. *Research in Race and Ethnic Relations*, 1, 127-167.
- Barlow, F. K., Paolini, S., Pedersen, A., Hornsey, M. J., Radke, H. R. M., Harwood, J., Rubin, M. (2012). The contact caveat: Negative contact predicts increased prejudice more than positive contact predicts reduced prejudice. *Personality and Social Psychology Bulletin*, 38(12), 1629-1643.
- Beelmann, A., & Heinemann, K. S. (2014). Preventing prejudice and improving intergroup attitudes: A meta-analysis of child and adolescent training programs. *Journal of Applied Developmental Psychology*, 35(1), 10-24. doi: <http://dx.doi.org/10.1016/j.appdev.2013.11.002>
- Ben-Moshe, D., & Halafoff, A. (2014). Antisemitism and Jewish children and youth in Australia's Capital Territory schools. *Social Inclusion*, 2(2), 47-56. doi: <http://www.cogitatiopress.com/ojs/index.php/socialinclusion/article/view/166>
- Block, K., Cross, S., Riggs, E., & Gibbs, L. (2014). Supporting schools to create an inclusive environment for refugee students. *International Journal of Inclusive Education*, 18, 1-19. doi: 10.1080/13603116.2014.899636
- Bolger, K. E., Patterson, C. J., & Kupersmidt, J. B. (1998). Peer relationships and self-esteem among children who have been maltreated. *Child Development*, 69(4), 1171-1197. doi: 10.1111/j.1467-8624.1998.tb06166.x
- Bosacki, S., Dane, A., Marini, Z., & YLC-CURA. (2007). Peer relationships and internalizing problems in adolescents: Mediating role of self-esteem. *Emotional and Behavioural Difficulties*, 12(4), 261-282. doi: 10.1080/13632750701664293
- Cardinia Shire Council. (2012). *Cardinia Shire population forecasts*. Retrieved from <http://forecast2.id.com.au/Default.aspx?id=213&pg=5000>.
- City of Casey Council. (2012). *City of Casey population forecasts*. Retrieved from <http://forecast2.id.com.au/Default.aspx?id=109&pg=5000>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences. (2nd ed.)*. New York, NY: Lawrence Erlbaum Associates.
- Cotton, K. (1993). Fostering intercultural harmony in schools: Research findings. *School Improvement Research Series*. Portland, OR: Office of Educational Research and Improvement
- Cumming, G., Fidler, F., Kalinowski, P., & Lai, J. (2012). The statistical

- recommendations of the American Psychological Association publication manual: Effect sizes, confidence intervals, and meta-analysis. *Australian Journal of Psychology*, 64(3), 138-146. doi: 10.1111/j.1742-9536.2011.00037.x
- Doyle, A. B., & Aboud, F. E. (1995). A longitudinal study of White children's racial prejudice as a social-cognitive development. *Merrill-Palmer Quarterly*, 41(2), 209-228.
- Dunn, K. M., & Geeraert, P. (2003). The geography of "race" and racisms. *GeoDate*, 16(3), 1-6.
- Farrell, P. (1995). *Children with emotional and behavioural difficulties: Strategies for assessment and intervention*. London, U.K.: Falmer Press.
- Field, A. (2009). *Discovering statistics using SPSS*. Thousand Oaks, CA: SAGE.
- Fischer, D. G., & Fick, C. (1993). Measuring social desirability: Short forms of the Marlowe-Crowne Social Desirability Scale. *Educational and Psychological Measurement*, 53(2), 417-424. doi: 10.1177/0013164493053002011
- Games, P. (1984). Data transformations, power, and skew: A rebuttal to Levine and Dunlap. *Psychological Bulletin*, 95(2), 345-347.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581-586.
- Greco, T., Priest, N., & Paradies, Y. (2010). *Review of strategies and resources to address race-based discrimination and support diversity in schools*. Carlton, VIC: Victorian Heath Promotion Foundation.
- Grigg, K., & Manderson, L. (2013). Monash University and Windermere Research 2013: Development and validation of an Australian measure of racial, ethnic, cultural, and religious acceptance. Melbourne, VIC: Windermere Child and Family Services.
- Grigg, K., & Manderson, L. (2014a). Developing the Australian Racism, Acceptance, and Cultural-Ethnocentrism Scale (RACES). Manuscript submitted for publication.
- Grigg, K., & Manderson, L. (2014b). Is there a relationship between psychopathic traits and racism? *Current Psychology*, Advanced online publication. doi: 10.1007/s12144-014-9283-9
- Grigg, K., & Manderson, L. (2014c). The Australian Racism, Acceptance, and Cultural-Ethnocentrism Scale (RACES): Item Response Theory findings. Manuscript submitted for publication.
- Gross, Z., & Rutland, S. D. (2014). Combatting antisemitism in the school playground: An Australian case study. *Patterns of Prejudice*, 48(3), 309-330. doi: 10.1080/0031322x.2014.918703
- Hickey, A., & Austin, J. (2009). *Building Racial Harmony: A report of project outcomes*. Toowoomba, QLD: University of Southern Queensland.
- Howarth, C. (2009). I hope we won't have to understand racism one day: Researching or reproducing race in social psychological research? *British Journal of Social Psychology*, 48(3), 407-426. doi: 10.1348/014466608x360727
- Janus, A. L. (2010). The influence of social desirability pressures on expressed immigration attitudes. *Social Science Quarterly*, 91(4), 928-946. doi: 10.1111/j.1540-6237.2010.00742.x
- Johnson, P. J., & Aboud, F. E. (2013). Modifying ethnic attitudes in young children: The impact of communicator race and message strength. *International Journal of Behavioral Development* (3), 182-191. doi: 10.1177/0165025412466522
- Lee, S. A., Gibbons, J. A., Thompson, J. M., & Timani, H. S. (2009). The Islamophobia Scale: Instrument development and initial validation. *International Journal for the Psychology of Religion*, 19(2), 92-105. doi: 10.1080/10508610802711137
- Loewenthal, K. M. (2001). *An introduction to psychological tests and scales (2nd ed.)*. London, U.K.: Psychology Press.
- Mullan, B., Todd, J., Chatzisarantis, N. L. D., & Hagger, M. S. (2014). Experimental methods in health psychology in Australia: Implications for applied

- research. *Australian Psychologist*, 49(2), 104-109. doi: 10.1111/ap.12046
- Neto, Y. F., & Pedersen, A. (2013). No time like the present: Determinants of intentions to engage in bystander anti-racism on behalf of Indigenous Australians. *Journal of Pacific Rim Psychology*, 7(1), 36-49. doi: doi:10.1017/prp.2013.4
- Nunnally, J. C. (1978). *Psychometric theory (2nd ed.)*. New York, NY: McGraw Hill.
- Ozdemir, S., Soydas, Y., & Grigg, K. (2012). *Building Harmony in the Cardinia Growth Corridor: Findings report 2011-2012 action research project*. Melbourne, VIC: Windermere Child and Family Services.
- Pachter, L. M., & Coll, C. G. (2009). Racism and child health: A review of the literature and future directions. *Journal of Developmental & Behavioral Pediatrics*, 30(3), 255-263.
- Pedersen, A., Beven, J., Walker, I., & Griffiths, B. (2004). Attitudes toward Indigenous Australians: The role of empathy and guilt. *Journal of Community & Applied Social Psychology*, 14(4), 233-249. doi: 10.1002/casp.771
- Pedersen, A., Walker, I., Paradies, Y., & Guerin, B. (2011). How to cook rice: A review of ingredients for teaching anti-prejudice. *Australian Psychologist*, 46, 55-63.
- Pedhazur, E. J., & Schmelkin, L., P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pettigrew, T. F., & Tropp, L. (2006). A meta-analytic test of Intergroup Contact Theory. *Journal of Personality and Social Psychology*, 90(5), 751-783.
- Phillips, S. T., & Ziller, R. C. (1997). Toward a theory and measure of the nature of nonprejudice. *Journal of Personality and Social Psychology*, 72(2), 420-434. doi: 10.1037/0022-3514.72.2.420
- Priest, N., Paradies, Y., Trenerry, B., Truong, M., Karlsen, S., & Kelly, Y. (2012). A systematic review of studies examining the relationship between reported racism and health and wellbeing for children and young people. *Social Science & Medicine*, 95, 115-127. doi: 10.1016/j.socscimed.2012.11.031
- Rutland, A., Cameron, L., Bennett, L., & Ferrell, J. (2005). Interracial contact and racial constancy: A multi-site study of racial intergroup bias in 3-5 year old Anglo-British children. *Journal of Applied Developmental Psychology*, 26(6), 699-713. doi: 10.1016/j.appdev.2005.08.005
- Salkind, N. (2006). *Exploring research (6th ed.)*. London, U.K.: Pearson Prentice Hall.
- Sanson, A., Augoustinos, M., Gridley, H., Kyrios, M., Reser, J., & Turner, C. (1998). Racism and prejudice: An Australian Psychological Society position paper. *Australian Psychologist*, 33(3), 161-182.
- Schneider, S. (2013). *Experimental design in the behavioral and social sciences*. London, U.K.: SAGE.
- Slomkowski, C., Klein, R. G., & Mannuzza, S. (1995). Is self-esteem an important outcome in hyperactive children? *Journal of Abnormal Child Psychology*, 23(3), 303-315. doi: 10.1007/bf01447559
- Stathi, S., Cameron, L., Hartley, B., & Bradford, S. (2014). Imagined contact as a prejudice-reduction intervention in schools: The underlying role of similarity and attitudes. *Journal of Applied Social Psychology*, 44(8), 536-546. doi: 10.1111/jasp.12245
- Strahan, R., & Gerbasi, K. C. (1972). Short, homogeneous versions of the Marlow-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 28(2), 191-193.
- Swinson, J. (2008). The self-esteem of pupils in schools for pupils with social, emotional and behavioural difficulties: Myth and reality. *British Journal of Special Education*, 35(3), 165-172. doi: 10.1111/j.1467-8578.2008.00389.x
- Tredoux, C. G., Noor, N. M., & de Paulo, L. (2009). Quantitative measures of respect and social inclusion in children: Overview and recommendations. *Effective Education*, 1(2), 169-186. doi: 10.1080/19415530903522568

- Tropp, L. (2005). Relationships between intergroup contact and prejudice among minority and majority status groups. *Psychological Science, 16*(12), 951-957.
- Tuffin, K. (2008). Racist discourse in New Zealand and Australia: Reviewing the last 20 years. *Social and Personality Psychology Compass, 2*(2), 591-607. doi: 10.1111/j.1751-9004.2007.00071.x
- Verkuyten, M., & Thijs, J. (2013). Multicultural education and inter-ethnic attitudes: An intergroup perspective. *European Psychologist, 18*(3), 179-190.
- Walker, I., & Crogan, M. (1998). Academic performance, prejudice, and the jigsaw classroom: New pieces to the puzzle. *Journal of Community & Applied Social Psychology, 8*(6), 381-393. doi: 10.1002/(sici)1099-1298(199811/12)8:6<381::aid-casp457>3.0.co2-6
- Wertheim, E. H., Davis, K., Freeman, E., & Trinder, M. (2010). Enhancing relationships in school communities: Creating respectful and cooperative relationships with the ERIS program. *Education Horizons, 11*(2), 21-23. doi: <http://pilambda.org/benefits/publications/educational-horizons/>
- White, F. A., & Abu-Raya, H. M. (2012). A Dual Identity-Electronic Contact (DIEC) experiment promoting short- and long-term intergroup harmony. *Journal of Experimental Social Psychology, 48*, 597-608. doi: 10.1016/j.jesp.2012.01.007
- White, F. A., Abu-Rayya, H. M., & Weitzel, C. (2014). Achieving twelve-months of intergroup bias reduction: The dual identity-electronic contact (DIEC) experiment. *International Journal of Intercultural Relations, 38*, 158-163. doi: <http://dx.doi.org/10.1016/j.ijintrel.2013.08.002>
- White, F. A., & Gleitzman, M. (2006). An examination of family socialisation processes as moderators of racial prejudice transmission between adolescents and their parents. *Journal of Family Studies, 12*(2), 247-260

Acknowledgements

The authors would like to acknowledge the support of Windermere Child and Family Services throughout the research. The authors would also like to acknowledge the assistance of Professor James Ogloff for his advice on preliminary versions of this article and for his general support throughout the research.

Address for Correspondence

Kaine Grigg
+61438600989
kaine.grigg@gmail.com