Effectiveness of School-Based Violence Prevention for Children and Youth

Cluster randomized controlled field trial of the Roots of Empathy program with replication and three-year follow-up

Robert G. Santos, Mariette J. Chartier, Jeanne C. Whalen, Dan Chateau and Leanne Boyd

Abstract

Aggression, bullying and violence in children and youth are prevalent in Canada (18%) and internationally. The authors evaluated the effectiveness of Roots of Empathy (ROE), a school-based mental health promotion and violence prevention program for children that has been widely implemented but rarely evaluated.

Eight school divisions were randomly assigned to either a treatment group that received ROE in 2002–2003 (445 students) or a wait-list control group (315 students). These were compared on three child mental health outcomes (physical aggression, indirect aggression and pro-social behaviour), rated by teachers and students (self-rated). The three wait-list school divisions received ROE in 2003–2004 (new cohort of 265 students) and were compared with the control group from 2002–2003 on the three outcomes, for replication purposes. For both comparisons, the authors report multi-level modelling analyses regarding (1) immediate effects after ROE completion at the end of the school year (pretest to post-test) and (2) long-term ROE effects up to three years after post-test.

ROE had replicated, beneficial effects on all teacher-rated outcomes, which were generally maintained or further improved across follow-up. However, ROE had almost no statistically significant or replicated effects on student-rated outcomes. This is the first evaluation to suggest that ROE appears effective when implemented on a large scale under real-world delivery conditions.

The health and well-being of Canada's children and youth, including their mental health, is a top priority for healthcare providers (Andresen 2006; Davidson 2011; Eggertson 2007; Kutcher 2011; McEwan et al. 2007; Sibbald 2006). Aggression, bullying and violence in children and youth are major public and population health problems internationally (Craig and Pepler 2003, 2007; Glew et al. 2005; Murray 2006; Nansel et al. 2001, 2004; Smith-Khuri et al. 2004; Williams et al. 2007), with Canada ranking dismally on bullying (26th) and victimization (27th) among 35 countries (Craig and Pepler 2007). Bullying is prevalent in Canada, where 18% of children have reported being bullied in the previous five days (Craig and Pepler 2003). Bullying is strongly associated with poor child physical health, mental health and psychosocial adjustment (Arsenault et al. 2006; Gini and Pozzoli 2009; Nansel et al. 2001, 2004; Pepler et al. 2011; Rigby 2003), including school and peer problems (Juvonen et al. 2003), youth violence (Nansel et al. 2003) and youth suicide (Kim and Leventhal 2008; Klomek et al. 2010). Herein, healthcare providers have four roles: identifying the problem, screening for mental health comorbidities, counselling families and advocating for violence prevention (American Medical Association Council on Scientific Affairs 2002; Pepler et al. 2011; Weir 2001; Wright 2005). For children with associated problems, the need for specialized treatment far exceeds the available supply (Davidson 2011; Kutcher 2011); developing and disseminating evidence-based mental health promotion and

bullying prevention approaches is therefore imperative (Waddell et al. 2005).

Improving children's social-cognitive skills can be efficacious in preventing chronic aggression and conduct problems (Dodge and Pettit 2003; Rutter et al. 1998). Because of their population reach, schools are a natural setting for mental health promotion and violence prevention (Kutcher 2011; Mytton et al. 2002; Patel et al. 2007) and classroom teaching is the most common efficacious approach (Cooper et al. 2000).

Because of their population reach, schools are a natural setting for mental health promotion and violence prevention.

Preventing childhood aggression, bullying and violence are top priorities for Canadian policy makers also. However, many remain critical of Canada's record at integrating research and practice in the prevention of child and youth mental health problems (Davidson 2011; Kutcher 2011; Kutcher and Davidson 2007; McLennan et al. 2004) because in widely disseminated programs, rigorous evaluation, or even minimal evidence of effectiveness, is frequently absent (Cooper et al. 2000; Smith et al. 2003). For example, several provincial governments have recently begun implementing Roots of Empathy (ROE) (Weir 2005), a new school-based, violence prevention program for children, developed in Canada (Gordon 2005) that has rarely been evaluated.

Intervention

Students in ROE participate in a structured, age-appropriate, 27-session curriculum (Gordon 2005) that is delivered to entire classrooms by trained, certified instructors. ROE centres on classroom visits by a family – a parent and his or her newborn infant. (There are nine pre-family sessions, nine family sessions and nine post-family sessions.) Therein, students observe parent-infant interaction and learn about early brain development, temperament, attachment, the reading of emotional cues, the conveyance of thoughts and feelings and social inclusion. ROE is based on theory that when children learn to label emotions and take the perspective of others, their empathy and pro-social behaviour increase, while their physical and indirect aggression decrease – thereby preventing violence (Gordon 2005).

As of the 2006–2007 school year, ROE was being implemented in over 2,000 kindergarten to grade eight classrooms across Canada, involving over 50,000 children and youth, with pilots in Australia, Japan and New Zealand (Schonert-Reichl and Hymel 2007). Similar numbers of students have been reached annually across Canada in subsequent school years through 2010–2011, with a cumulative estimated total of 363,000 students reached since 1996, according to the ROE website (2011). In 2008, the Assembly of First Nations passed a resolution endorsing ROE, describing it as "compatible with traditional First Nations' teachings and worldviews." Notwithstanding this widespread use and support, to date there has been no published peer-reviewed evaluation of ROE. Most evaluated school-based violence prevention programs have been models or demonstrations that were evaluated for efficacy only; their real-world effectiveness is largely unknown (Wilson et al. 2003; Wilson and Lipsey 2007).

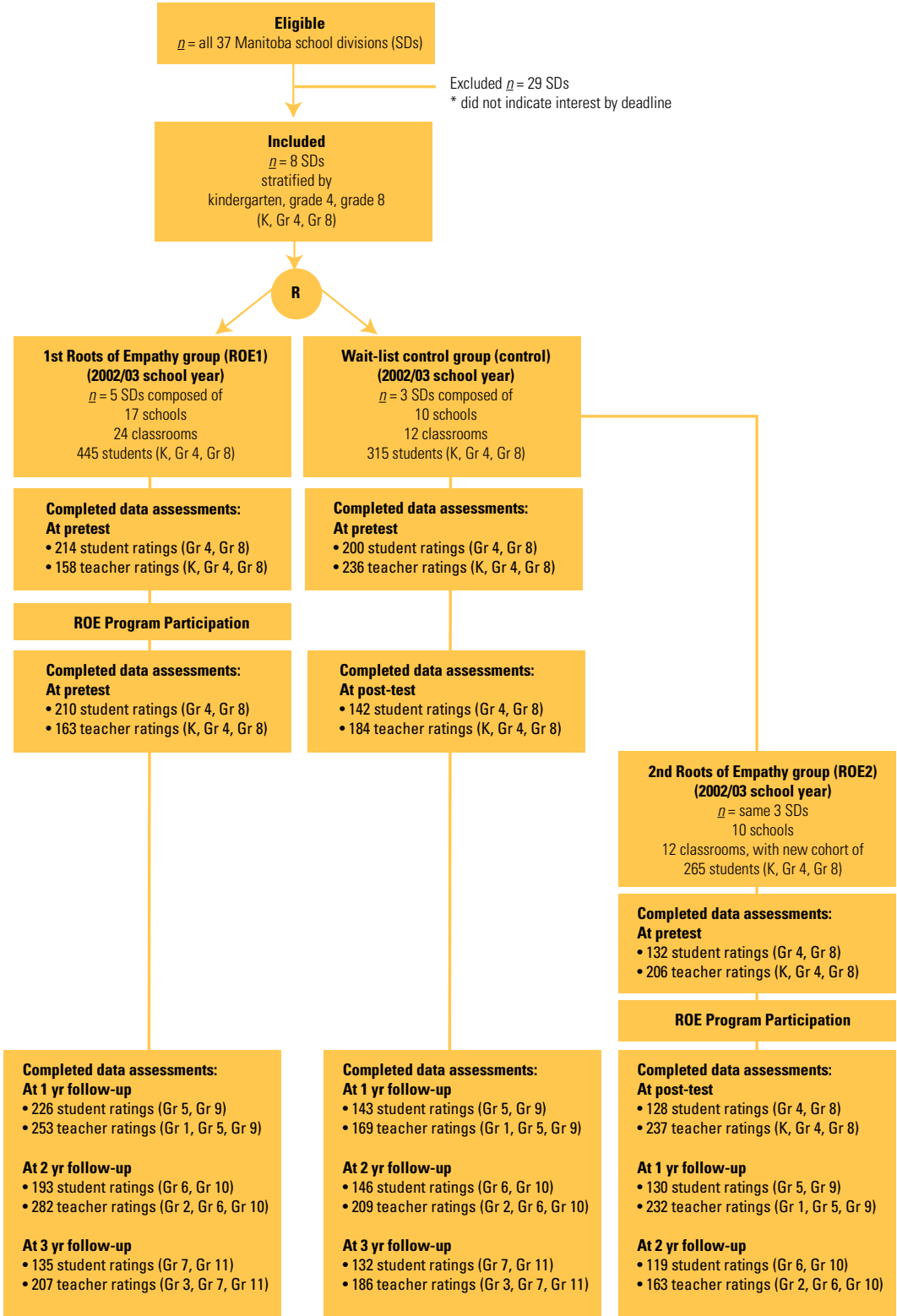
In this article, we follow the Consolidated Standards of Reporting Trials (CONSORT) guidelines (Campbell et al. 2004). The objective of our study was to evaluate the real-world effectiveness of ROE in preventing violence (reducing aggression and increasing pro-social behaviour) in children and youth at the individual level, immediately after program completion and up to three years afterwards, in two successive samples determined via cluster random assignment, in order to provide rigorous evidence to inform provincial government decision-making regarding the future expansion of ROE in Manitoba, including questions related to relative effectiveness by student gender and grade level.

Methods

In 2002, a "natural experiment" opportunity arose in Manitoba to rigorously evaluate ROE. Limited provincial government funding was available to implement ROE in five school divisions in the 2002–2003 school year. In June 2002, the provincial government's Healthy Child Manitoba Office (HCMO) invited all 37 public school divisions to express their interest in ROE. Eight school divisions expressed interest by the August 30, 2002, deadline and were eligible for ROE funding. All eight school divisions agreed with our proposal that cluster random assignment (at the school division level) was the fairest approach to ROE resource allocation. This design also reduced the likelihood of ROE "spillover effects" between treatment and control groups (Campbell et al. 2004). School divisions also agreed with our proposal to pre-stratify implementation along three grades (kindergarten, grade four and grade eight) to examine the relative effectiveness of ROE for different grades. School divisions prioritized and identified classrooms in each of these three grades for ROE implementation prior to random assignment in September 2002 and government-funded ROE training in October 2002.

As a quality assurance study, this ROE program evaluation did not require Research Ethics Board review (Canadian Institutes of Health Research et al. 1998, 2010). It was not registered as a clinical trial. Under the *The Healthy Child Manitoba Act*, HCMO is legislatively mandated to evaluate provincial government programs for children and youth.

FIGURE 1.
Selection and flow of clusters and individual participants through the phases of the randomized trial



Randomization

As illustrated in Figure 1, in this cluster randomized controlled field trial, HCMO randomly assigned the eight school divisions to either a treatment group that received ROE in the 2002–2003 school year (ROE1; 445 students) or a wait-list control group (315 students). HCMO used a computerized random number generator for the random assignment process. Sample sizes were determined by the number of students in each of the classrooms prioritized and selected by school divisions prior to randomization.

Measurement

In October 2002, following written school division notification to parents regarding the ROE evaluation, HCMO collected sociodemographic data (student gender and grade level) and pretested ROE1 and control groups on three child mental health outcomes: physical aggression (6 items: e.g., threatening people, bullying others, kicking or hitting other children), indirect aggression (5 items: e.g., trying to get others to dislike a person, telling a person's secrets to a third person) and pro-social behaviour (10 items: e.g., comforting a child who is crying or upset, offering to help other children who are having difficulty, inviting others to join a game). These were rated by teachers (kindergarten, grade four, grade eight) and self-rated by students (grade four, grade eight) using parallel instruments previously validated in Canada's National Longitudinal Survey of Children and Youth (Human Resources Development Canada and Statistics Canada 1996). These individual-level instruments served as our primary outcome measures of violence prevention.

Using the same three measures, rated by teachers and students, we post-tested ROE1 and control groups at the end of the 2002–2003 school year and annually for three years thereafter. Neither students nor teachers were blinded to group assignment at pretest or post-test. It is unlikely that the teachers who provided the three annual follow-up ratings were aware of group assignment (i.e., which children in their class had previously participated in ROE or not), but we did not measure this awareness directly.

The three school divisions randomized to the wait list received ROE in the subsequent 2003–2004 school year (ROE2; new cohort of 265 students). ROE2 was pretested and

TABLE 1.

Physical aggression: observed (unadjusted) mean scores and SDs for ROE1, control group and ROE2, across measurement time points, as rated by teachers and students (range of score: 0–12)

			Teacher Ratings		
Comparison Group	**Pretest**	**Post-Test**	**1 y Follow-Up***	**2 y Follow-Up***	**3 y Follow-Up***
ROE1					
Mean score	2.06	1.52	1.71	1.21	1.10
SD	2.87	2.70	1.78	2.24	1.80
n	158	160	261	288	206
Control group					
Mean score	1.10	1.56	2.14	1.49	1.44
SD	2.22	2.82	2.11	2.60	2.58
n	243	193	169	215	188
ROE2					
Mean score	1.61	1.52	1.92	1.68	
SD	2.67	2.73	2.75	3.08	
n	212	240	233	168	
			Student Ratings		
Comparison Group	**Pretest**	**Post-Test**	**1 y Follow-Up***	**2 y Follow-Up***	**3 y Follow-Up***
ROE1					
Mean score	2.13	2.26	2.69	1.66	2.19
SD	2.07	2.10	1.73	1.88	2.15
n	219	208	234	195	134
Control group					
Mean score	2.01	1.88	2.76	1.63	1.80
SD	2.41	2.28	1.79	1.87	1.97
n	200	146	145	151	137
ROE2					
Mean score	2.18	2.11	2.25	1.80	
SD	1.94	2.22	2.20	1.98	
n	140	139	129	117	

*Follow-up data were collected annually for three years for ROE1 and collected annually for two years for ROE2.

ROE1 = first Roots of Empathy group; ROE2 = second Roots of Empathy group; SD = standard deviation.

TABLE 2.
Indirect aggression: observed (unadjusted) mean scores and SDs for ROE1, control group and ROE2, across measurement time points, as rated by teachers and students (range of score: 0–10)

			Teacher Ratings		
Comparison Group	**Pretest**	**Post-Test**	**1 y Follow-Up***	**2 y Follow-Up***	**3 y Follow-Up***
ROE1					
Mean score	2.35	1.53	1.27	1.66	1.25
SD	2.89	2.15	2.15	2.53	2.11
n	159	162	261	264	202
Control group					
Mean score	1.41	2.02	1.68	1.33	1.49
SD	2.25	2.66	2.56	2.21	2.35
n	235	188	170	212	186
ROE2					
Mean score	1.61	1.56	1.90	1.64	
SD	2.34	2.27	2.42	2.81	
n	208	239	225	161	
			Student Ratings		
Comparison Group	**Pretest**	**Post-Test**	**1 y Follow-Up***	**2 y Follow-Up***	**3 y Follow-Up***
ROE1					
Mean score	2.10	1.87	1.68	1.35	1.49
SD	2.01	1.76	1.92	1.62	1.66
n	219	209	235	195	136
Control group					
Mean score	2.16	2.21	1.88	1.33	1.79
SD	2.19	2.10	2.13	1.58	1.95
n	204	148	147	157	133
ROE2					
Mean score	1.77	1.81	1.78	1.54	
SD	1.85	1.86	1.89	1.89	
n	141	135	140	118	

*Follow-up data were collected annually for three years for ROE1 and collected annually for two years for ROE2.
ROE1 = first Roots of Empathy group; ROE2 = second Roots of Empathy group; SD = standard deviation.

post-tested in 2003–2004 and followed up for two years afterwards, on all outcomes, and compared with the control group from the 2002–2003 sample (see Figure 1).

Analyses

Given clustering in data and ROE delivery, we used multi-level modelling (SAS PROC MIXED) to account for three levels of variability: intra-individual change in students over time (in the three outcomes), inter-individual differences between students (gender) and inter-group differences between classrooms (assignment to ROE or control group, grade level). The latter level encompassed the school division level of randomization, as well as the school level, given that these were fixed (i.e., stratified by grade level and identified as blocks for inclusion in the evaluation prior to randomization). Intra-class correlation coefficients (ICCs: variance between classes divided by [variance between classes + variance within classes + residual]) at pretest (range: 0.11-0.29) indicated that a considerable amount of variance was due to variation between classrooms (i.e., students in the same classroom were more similar to each other than to students from other classrooms). We found ICCs similar to those in other school-based studies (Bloom et al. 2007; Hedges and Hedberg 2007; Raudenbush et al. 2007). Multi-level modelling accounted for clustering within classrooms and over time and incorporated all participants who were observed at least once (Allison 2002; Donner and Klar 2004; Murray et al. 2004; Raudenbush and Bryk 2002; Schafer and Graham 2002).

Overall, HCMO obtained data on 93% of the original sample, with approximately 50% of observations missing at any time point (see Figure 1). Our multi-level modelling used maximum likelihood estimation, allowing for results to be interpreted as if there were no missing data, under the assumption that data were missing at random (Allison 2002; Raudenbush and Bryk 2002; Schafer and Graham 2002).

For both comparisons (ROE1 versus control and ROE2 versus control), we report multi-level modelling (intention to treat) analyses for the three outcomes regarding (1) immediate effects after ROE completion at school year end (pretest to post-test) and (2) long-term ROE effects over follow-up (post-test through one year, two years and up to three years). We also

explored student gender and grade level as potential moderators of ROE effects.

Results

For each outcome, Tables 1, 2 and 3 show observed (unadjusted) scores for all groups across measurement times, as rated by teachers and students.

At pretest, teacher ratings of ROE1 and ROE2 indicated statistically significantly higher physical aggression (ROE1 only), higher indirect aggression and lower pro-social behaviour compared with the control group (all $p < .05$). However, student ratings were generally similar between groups (except higher pro-social behaviour in ROE2, $p < .05$).

We found acceptable levels of internal consistency reliability (Cronbach's alpha) for all outcomes across all groups and measurement times, with generally higher coefficients for teacher ratings (range .80–.95) than student ratings (range .67–.94).

As shown in Table 4, multilevel analyses found that, as rated by teachers, ROE had beneficial immediate effects on all outcomes, reducing physical aggression and indirect aggression and increasing prosocial behaviour, replicated in both ROE1 and ROE2. By comparison, as self-rated by students, ROE effects were less pronounced and fewer were statistically significant or replicated. For all outcomes across measurement times, teacher ratings and student self-ratings were not highly correlated (mean Pearson rs of 0.30, 0.20, and 0.28 for physical aggression, indirect aggression, and prosocial behaviour, respectively).

As shown in Table 5, multilevel analyses found that, as rated by teachers, beneficial outcomes were generally maintained (as indicated by the lack of statistically significant differences between groups) or continued to improve following ROE completion, with one exception: Some of the ROE1 gain in prosocial behaviour was not maintained.

TABLE 3.
Pro-social behaviour: observed (unadjusted) mean scores and SDs for ROE1, control group and ROE2, across measurement time points, as rated by teachers and students (range of score: 0–20)

			Teacher Ratings		
Comparison Group	**Pretest**	**Post-Test**	**1 y Follow-Up***	**2 y Follow-Up***	**3 y Follow-Up***
ROE1					
Mean score	8.84	11.23	10.36	10.61	10.23
SD	5.57	5.57	5.91	5.52	5.12
n	158	163	253	282	207
Control group					
Mean score	10.68	11.28	10.09	9.67	9.49
SD	4.95	4.89	5.07	5.90	5.37
n	236	184	169	209	186
ROE2					
Mean score	9.37	11.21	9.21	10.35	
SD	5.24	5.10	5.45	5.99	
n	206	237	232	163	
			Student Ratings		
Comparison Group	**Pretest**	**Post-Test**	**1 y Follow-Up***	**2 y Follow-Up***	**3 y Follow-Up***
ROE1					
Mean score	12.44	12.46	12.86	13.08	12.37
SD	3.99	3.78	3.87	3.30	3.37
n	214	210	226	193	135
Control group					
Mean score	12.23	12.27	12.52	13.03	12.84
SD	4.09	4.03	3.61	3.41	3.47
n	200	142	143	146	132
ROE2					
Mean score	13.31	13.77	12.62	13.51	
SD	3.78	3.55	3.65	3.19	
n	132	128	130	119	

*Follow-up data were collected annually for three years for ROE1 and collected annually for two years for ROE2.
ROE1 = first Roots of Empathy group; ROE2 = second Roots of Empathy group; SD = standard deviation.

Moderators of ROE

Most interactions between ROE and student gender or grade level were inconsistent across samples, but multi-level analyses may suggest that (1) immediately after completion, ROE is more effective in decreasing indirect aggression in girls than in boys and in improving pro-social behaviour for younger students than for older students; and (2) over follow-up, gains in pro-social behaviour may fade in boys or in older children. All of these interaction effects are exploratory and require replication.

TABLE 4

Immediate effects of the Roots of Empathy (ROE) program in violence prevention: Results of multilevel modeling analyses from pretest to posttest, in ROE1-control group and ROE2-control group comparisons, as rated by teachers and students

	ROE1-control group comparison			ROE2-control group comparison		
Child mental health outcomes	**Regression estimate (95% CI)**	**Effect size[a] (95% CI)**	***p***	**Regression estimate (95% CI)**	**Effect size[a] (95% CI)**	***p***
Teacher-rated:						
Physical aggression	-0.64 (-1.09 – -0.20)	-0.25 (-0.43 – -0.08)	0.01*	-0.38 (-0.71 – 0.05)	-0.15 (-0.28 – -0.02)	0.03*
Indirect aggression	-0.30 (-1.81 – -0.80)	-0.51 (-0.70 – -0.31)	0.00*	-0.66 (-1.08 – -0.24)	-0.26 (-0.42 – -0.10)	0.00*
Prosocial behaviour	1.08 (0.43 – 2.12)	0.21 (0.01 – 0.40)	0.04*	0.97 (0.08 – 1.86)	0.18 (0.14 – 0.35)	0.04*
Student-rated:						
Physical aggression	0.18 (-0.29 – 0.65)	0.08 (-0.13 – 0.29)	0.45	0.06 (-0.47 – 0.60)	0.03 (-0.21 – 0.27)	0.82
Indirect aggression	-0.41 (-0.86 – 0.03)	-0.20 (-0.41 – -0.02)	0.07	-0.06 (-0.53 – 0.41)	-0.03 (-0.25 – 0.20)	0.80
Prosocial behaviour	0.53 (-0.20 – 1.26)	0.13 (-0.05 – 0.31)	0.15	0.88 (0.04 – 1.71)	0.22 (0.01 – 0.42)	0.04*

[a]Effect size is calculated using the following formula: unstandardized regression estimate divided by the pooled standard deviation of the outcome of the sample. (Hedges, 2007)

* Difference between program group and control group is statistically significant ($p < .05$).

CI = confidence interval. ROE1 = first Roots of Empathy group. ROE2 = second Roots of Empathy group.

TABLE 5.

Long-term effects of the Roots of Empathy (ROE) program in violence prevention: Results of multilevel modeling analyses from posttest through up to 3 years follow-up after program completion, in ROE1-control group and ROE2-control group comparisons, as rated by teachers and students

	ROE1-control group comparison			ROE2-control group comparison		
Child mental health outcomes	**Regression estimate (95% CI)**	**Effect size[a] (95% CI)**	***p***	**Regression estimate (95% CI)**	**Effect size[a] (95% CI)**	***p***
Teacher-rated:						
Physical Aggression	-0.15 (-0.23 – -0.07)	-0.06 (-0.09 – -0.03)	0.00*	0.34 (-0.34 – 1.01)	0.14 (-0.14 – 0.43)	0.33
Indirect Aggression	-0.06 (-0.16 – 0.03)	-0.03 (-0.06 – 0.01)	0.18	-0.04 (-0.33 – 0.24)	-0.02 (-0.14 – 0.10)	0.76
Prosocial Behaviour	-0.65 (-0.88 – -0.43)	-0.12 (-0.17 – -0.08)	0.00*	-0.20 (-0.47 – 0.08)	-0.08 (-0.20 – 0.03)	0.17
Student-rated:						
Physical Aggression	-0.00 (-0.09 – 0.08)	-0.00 (-0.04 – 0.04)	0.99	-0.10 (-0.36 – 0.16)	-0.04 (-0.15 – 0.07)	0.46
Indirect Aggression	-0.05 (-0.13 – 0.04)	-0.02 (-0.06 – 0.02)	0.26	0.14 (-0.10 – 0.38)	0.06 (-0.04 – 0.16)	0.25
Prosocial Behaviour	-0.19 (-1.63 – 1.26)	-0.08 (-0.69 – 0.53)	0.01*	0.23 (-0.14 – 0.60)	0.10 (-0.06 – 0.25)	0.23

a Effect size is calculated using the following formula: unstandardized regression estimate divided by the pooled standard deviation of the outcome of the sample. (Hedges, 2007)

* Difference between program group and control group is statistically significant ($p < .05$).

CI = confidence interval. ROE1 = first Roots of Empathy group. ROE2 = second Roots of Empathy group.

Discussion and Conclusion

This is the first evaluation to suggest that ROE is effective when implemented on a large scale under real-world conditions. On average, ROE seems to achieve replicable immediate effects, as rated by teachers (mean absolute effect size [ES] = .25), larger effects than those reported in the most comprehensive meta-analysis to date regarding similar programs implemented as smaller-scale models or demonstrations (mean ES = .21) or in routine practice (mean ES = .10) (Wilson et al. 2003; Wilson and Lipsey 2007). Compared with other systematic reviews, our results show that ROE appears to be as effective as, or more effective than, similar programs that have targeted high-risk students (Mytton et al. 2002) or employed curricula, school-wide approaches or social skills training (Vreeman and Carroll 2007). Translated into everyday terms, if an estimated 15% of schoolchildren get into a fight in a school year, an ES of 0.25 for ROE represents a reduction in fighting to about 8%, approximately half the baseline rate (Wilson et al. 2003). This suggests practical significance and, given associated morbidity, probable clinical importance from the perspectives of mental health promotion and mental illness prevention. The enhancement of empathy and the promotion of optimal social contact are also essential to reducing mental health stigma (Hinshaw and Stier 2008; Stuart et al. 2011). At an estimated cost for ROE of C$108 per child per year (C$4 per child per session for 27 sessions), ROE has high potential cost-effectiveness given the enormous cost of conduct disorder alone (an estimated average of C$7,944 per child per year from age 10 to 28; Scott et al. 2001). A second cluster randomized trial of ROE, in British Columbia (Schonert-Reichl et al. 2007, March), appears to replicate our immediate effects; it also plans a three-year follow-up. ROE appears close to meeting international standards of evidence for effectiveness (Flay et al. 2005). Few studies of similar programs have followed long-term effects. Our findings suggest that ROE may be beneficial up to three years after completion.

While we found similar results in ROE1 and ROE2, ESs in the latter were more modest. This may be attributable to (1) ROE2 and the control group being from the same school divisions; (2) larger pretest aggression scores in ROE1 (previous studies found larger pretest aggression scores predict larger program ESs; Wilson and Lipsey 2007); (3) differential quality of implementation (Wilson and Lipsey 2007); or (4) school or community context (Hughes et al. 2005; Metropolitan Area Child Study Group 2007). We did not measure implementation or context, and these merit future measurement.

Student self-rated ROE effects were smaller than teacher-rated effects, and fewer were statistically significant or replicated. Evaluations of similar programs typically find smaller ESs when using student ratings, and only 22% of studies used them (Wilson and Lipsey 2007). Our modest correlations between student and teacher ratings are consistent with the psychopathology literature; each informant may contribute different but useful information (Achenbach et al. 2005; De Los Reyes and Kazdin 2005).

Strengths of our evaluation include internal validity and ecological validity through a rigorous design (cluster random assignment with multiple outcome informants and longitudinal follow-up) to evaluate ROE under real-world conditions. Our results are notable as many efficacious interventions do not improve outcomes when exported from laboratory conditions into routine practice (Flay et al. 2005), particularly in the first implementation year. We encourage other governments to evaluate untested programs via random assignment prior to larger-scale implementation.

We encourage other governments to evaluate untested programs via random assignment prior to larger-scale implementation.

Our evaluation had limitations. As a natural experiment, our evaluation did not calculate a priori statistical power (Guittet et al. 2005). Our statistically significant findings suggest that our ICCs and number of groups per condition (the two primary determinants of power in cluster randomized trials; Murray et al. 2004) permitted sufficient statistical power. By chance, ROE1 and ROE2 differed from the control group at pretest and, due to limited resources, observations for some time points were missing. Both were addressed through multi-level modelling, which controlled for pretest differences (and other potential confounders over time) and provided robust maximum likelihood estimates of missing data.

Canada faces continuing challenges in improving child and youth mental health, particularly in prevention (Andresen 2006; Davidson 2011; Eggertson 2005; Kutcher 2011; McLennan et al. 2004; Waddell et al. 2005, 2007). As with other major public health problems, the burden of suffering associated with aggression, bullying and violence will not be significantly reduced by clinical services alone; effective prevention programs are also urgently needed (Craig and Pepler 2003; Offord et al. 1998; Waddell et al. 2005, 2007). Our evaluation suggests that ROE is effective and worthy of consideration in emerging evidence-based mental health strategies for children and youth across Canada. HQ

Acknowledgement

The views expressed herein are those of the authors and do not necessarily reflect the views of the government of Manitoba.

References

Achenbach, T.M., R.A. Krukowski, L. Dumenci and M.Y. Ivanova. 2005. "Assessment of Adult Psychopathology: Meta-analyses and

Implications of Cross-informant Correlations." *Psychological Bulletin* 131(3): 361–82.

Allison, P.D. 2002. *Missing Data.* Thousand Oaks, CA: Sage.

American Medical Association Council on Scientific Affairs. 2002. *Bullying Behaviors among Children and Adolescents* (CSA Rep. No. 1, A-02). Chicago: American Medical Association. Retrieved January 21, 2011. <http://www.ama-assn.org/ama1/pub/upload/mm/443/csaa-02.pdf>.

Andresen, M. 2006. "Mental Health Moves Up the Agenda." *Canadian Medical Association Journal* 175(2): 139–40.

Arseneault, L., E. Walsh, K. Trzesniewski, R. Newcombe, A. Caspi and T.E. Moffitt. 2006. "Bullying Victimization Uniquely Contributes to Adjustment Problems in Young Children: A Nationally Representative Cohort Study." *Pediatrics* 118(1): 130–38.

Assembly of First Nations. 2008. *Resolution 38: Support for Roots of Empathy* (Annual General Assembly, July 2008, Quebec City, QC) Ottawa, ON: Author. <www.afn.ca/uploads/files/aga-res-08.pdf>.

Bloom, H.S., L. Richburg-Hayes and A.R. Black. 2007. "Using Covariates to Improve Precision for Studies that Randomize Schools to Evaluate Educational Interventions." *Educational Evaluation and Policy Analysis* 29(1): 30–59.

Campbell, M.K., D.R. Elbourne and D.G. Altman for the CONSORT Group. 2004. "CONSORT Statement: Extension to Cluster Randomised Trials." *BMJ* 328: 702–08.

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada and Social Sciences and Humanities Research Council of Canada. 1998. *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* [with 2000, 2002, 2005 amendments]. Ottawa, ON: Interagency Secretariat on Research Ethics.

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada and Social Sciences and Humanities Research Council of Canada. 2010. *Tri-council Policy Statement: Ethical Conduct for Research Involving Humans* (2nd ed.). Ottawa, ON: Interagency Advisory Panel on Research Ethics.

Cooper, W.O., M. Lutenbacher and K. Faccia. 2000. "Components of Effective Youth Violence Prevention Programs for 7- to 14-Year-Olds." *Archives of Pediatrics and Adolescent Medicine* 154: 1134–39.

Craig, W.M. and D.J. Pepler. 2003. "Identifying and Targeting Risk for Involvement in Bullying and Victimization." Canadian Journal of Psychiatry 48(9): 577–82.

Craig, W.M. and D.J. Pepler. 2007. "Understanding Bullying: From Research to Practice." *Canadian Psychology* 48(2): 86–95.

Davidson, S. 2011. "The State of Child and Youth Mental Health in Canada: Past Problems and Future Fantasies." *Healthcare Quarterly* 14(Special Issue): 9–13.

De Los Reyes, A. and A.E. Kazdin. 2005. "Informant Discrepancies in the Assessment of Childhood Psychopathology: A Critical Review, Theoretical Framework, and Recommendations for Further Study." *Psychological Bulletin* 131(4): 483–509.

Dodge, K.A. and G.S. Pettit. 2003. "A Biopsychosocial Model of the Development of Chronic Conduct Problems in Adolescence." *Developmental Psychology* 39(2): 349–71.

Donner, A. and N. Klar. 2004. "Pitfalls of and Controversies in Cluster Randomization Trials." *American Journal of Public Health* 94(3): 416–22.

Eggertson, L. 2005. "Children's Mental Health Services Neglected: Kirby." *Canadian Medical Association Journal* 173(5): 471.

Eggertson, L. 2007. "Physicians Challenge Canada to Make Children, Youth a Priority." *Canadian Medical Association Journal* 176(12): 1693–94.

Flay, B.R., A. Biglan, R.F. Boruch, F.G. Castro, D. Gottfredson, S. Kellam et al. 2005. "Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination." *Prevention Science* 6(3): 151–75.

Gini, G. and T. Pozzoli. 2009. "Association between Bullying and Psychosomatic Problems: A Meta-analysis." *Pediatrics* 123(3): 1059–65.

Glew, G.M., M. Fan, W. Katon, F.P. Rivara and M.A. Kernic. 2005. "Bullying, Psychosocial Adjustment, and Academic Performance in Elementary School." *Archives of Pediatrics and Adolescent Medicine* 159: 1026–31.

Gordon, M. 2005. *Roots of Empathy: Changing the World Child by Child.* Toronto, ON: Thomas Allen.

Guittet, L., B. Girardeau and P. Ravaud. 2005. "A Priori Postulated and Real Power in Cluster Randomized Trials: Mind the Gap." *BMC Medical Research Methodology* 5:25. Retrieved January 21, 2011. <www.biomedcentral.com/1471-2288/5/25>.

Hedges, L.V. 2007. "Effect Sizes in Cluster-Randomized Designs." *Journal of Educational and Behavioral Statistics* 32: 341-70.

Hedges, L.V. and E.C. Hedberg. 2007. "Intraclass Correlation Values for Planning Group-Randomized Trials in Education." *Educational Evaluation and Policy Analysis* 29(1): 60–87.

Hinshaw, S.P. and A. Stier. 2008. "Stigma as Related to Mental Disorder." *Annual Review of Clinical Psychology* 4: 367–93.

Hughes, J.N., T.A. Cavell, B.T. Meehan, D. Zhang and C. Collie. 2005. "Adverse School Context Moderates the Outcomes of Selective Intervention for Aggressive Children." *Journal of Consulting and Clinical Psychology* 73(4): 731–36.

Human Resources Development Canada and Statistics Canada. 1996. *Growing Up in Canada: National Longitudinal Survey of Children and Youth.* Ottawa, ON: Statistics Canada.

Juvonen, J., S. Graham and M.A. Schuster. 2003. "Bullying among Young Adolescents: The Strong, the Weak, and the Troubled." *Pediatrics* 112(6): 1231–37.

Kim, Y.S. and B.L. Leventhal. 2008. "Bullying and Suicide: A Review." *International Journal of Adolescent Medicine and Health* 20(2): 133–54.

Klomek, A.B., A. Sourander and M. Gould. 2010. "The Association of Suicide and Bullying in Childhood: A Review of Cross-Sectional and Longitudinal Research Findings." *Canadian Journal of Psychiatry* 55(5): 282–88.

Kutcher, S. 2011. "Facing the Challenge of Care for Child and Youth Mental Health in Canada: A Critical Commentary, Five Suggestions for Change and a Call to Action." *Healthcare Quarterly* 14(Special Issue): 14–22.

Kutcher, S. and S. Davidson. 2007. "Mentally Ill Youth: Meeting Service Needs." *Canadian Medical Association Journal* 176(4): 417.

McEwan, K., C. Waddell and J. Barker. 2007. "Bringing Children's Mental Health 'Out of the Shadows.'" *Canadian Medical Association Journal* 176(4): 471–72.

McLennan, J.D., H.L. MacMillan and E. Jamieson. 2004. "Canada's Programs to Prevent Mental Health Problems in Children: The Research-Practice Gap." *Canadian Medical Association Journal* 171(9): 1069–71.

Metropolitan Area Child Study Research Group. 2007. "Changing the Way Children 'Think' about Aggression: Social Cognitive Effect of a Preventive Intervention." *Journal of Consulting and Clinical Psychology*

75(1): 160–67.

Murray, D.M., S.P. Varnell and J.L. Blitstein. 2004. "Design and Analysis of Group-Randomized Trials: A Review of Recent Methodological Developments." *American Journal of Public Health* 94(3): 423–32.

Murray, S. 2006. "Public Health." *Canadian Medical Association Journal* 174(5): 620–21.

Mytton, J.A., C. DiGuiseppi, D.A. Gough, R.S. Taylor and S. Logan. 2002. "School-Based Violence Prevention Programs: Systematic Review of Secondary Prevention Trials." *Archives of Pediatrics and Adolescent Medicine* 156: 752–62.

Nansel, T.R., W. Craig, M.D. Overpeck, G. Saluja, W.J. Ruan and the Health Behaviour in School-Aged Children Bullying Analyses Working Group. 2004. "Cross-National Consistency in the Relationship between Bullying Behaviors and Psychosocial Adjustment." *Archives* of Pediatrics and Adolescent Medicine 158: 730–36.

Nansel, T.R., M. Overpeck, D.L. Haynie, W.J. Ruan and P.C. Scheidt. 2003. "Relationships between Bullying and Violence among US Youth." *Archives of Pediatrics and Adolescent Medicine* 157: 348–53.

Nansel, T.R., M. Overpeck, R.S. Pilla, W.J. Ruan, B. Simons-Morton and P. Scheidt. 2001. "Bullying Behaviors among US Youth: Prevalence and Association with Psychosocial Adjustment." *Journal of the American Medical Association* 285(16): 2094–100.

Offord, D.R., H.C. Kraemer, A.E. Kazdin, P.S. Jensen and R. Harrington. 1998. "Lowering the Burden of Suffering from Child Psychiatric Disorder: Trade-Offs among Clinical, Targeted, and Universal Interventions." *Journal of the American Academy of Child and Adolescent Psychiatry* 37(7): 686–94.

Patel, V., A.J. Flisher, S. Hetrick and P. McGorry. 2007. "Mental Health of Young People: A Global Public-Health Challenge." *Lancet* 369: 1302–13.

Pepler, D.J., J. German, W. Craig and S. Yamada. 2011. "Why Worry about Bullying?" *Healthcare Quarterly* 14(Special Issue): 72–79.

Raudenbush, S.W., A. Martinez and J. Spybrook. 2007. "Strategies for Improving Precision in Group-Randomized Experiments." *Educational Evaluation and Policy Analysis* 29(1): 5–29.

Raudenbush, S.W. and A.S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods* (2nd ed.). Thousand Oaks, CA: Sage.

Rigby, K. 2003. "Consequences of Bullying in Schools." *Canadian Journal of Psychiatry* 48(9): 583–90.

Roots of Empathy. 2011. *Roots of Empathy in Canada.* Toronto, ON: Author. Retrieved March 1, 2011. <http://www.rootsofempathy.org/en/where-we-are/canada.html>.

Rutter, M., H. Giller and A. Hagell. 1998. *Antisocial Behavior by Young People.* New York, NY: Cambridge University Press.

Schafer, J.L and J.W. Graham. 2002. "Missing Data: Our View of the State of the Art." *Psychological Methods* 7(2): 147–77.

Schonert-Reichl, K.A. and S. Hymel. 2007. "Educating the Heart as Well as the Mind: Social and Emotional Learning for School and Life Success." *Education Canada* 47(2): 20–25.

Schonert-Reichl, K.A., V. Smith and C. Hertzman. 2007, March. *Promoting Emotional Competence in School-Aged Children: An Experimental Trial of the 'Roots of Empathy' Program.* Poster presented at Biennial Meeting of the Society for Research in Child Development, Boston, MA.

Scott, S., M. Knapp, J. Henderson and B. Maughan. 2001. "Financial Cost of Social Exclusion: Follow-Up Study of Antisocial Children into Adulthood." *BMJ* 323: 191–95.

Sibbald, B. 2006. "CMA Wants National Children's Health Strategy." *Canadian Medical Association Journal* 175(6): 567.

Smith, P.K., K. Ananiadou and H. Cowie. 2003. "Interventions to Reduce School Bullying." *Canadian Journal of Psychiatry* 48(9): 591–99.

Smith-Khuri, E., R. Iachan, P.C. Scheidt, M.D. Overpeck, S.N. Gabhainn, W. Pickett et al. 2004. "A Cross-National Study of Violence-Related Behaviors in Adolescents." *Archives of Pediatrics and Adolescent Medicine* 158: 539–44.

Stuart, H., M. Koller, R. Christie and M. Pietrus. 2011. "Reducing Mental Health Stigma: A Case Study." *Healthcare Quarterly* 14(Special Issue): 40–49.

Vreeman, R.C. and A.E. Carroll. 2007. "A Systematic Review of School-Based Interventions to Prevent Bullying." *Archives of Pediatrics and Adolescent Medicine* 161: 78–88.

Waddell, C., K. McEwan, C.A. Shepherd, D.R. Offord and J.M. Hua. 2005. "A Public Health Strategy to Improve the Mental Health of Canadian Children." *Canadian Journal of Psychiatry* 50(4): 226–33.

Waddell, C., K. McEwan, R.D. Peters, J.M. Hua and O. Garland. 2007. "Preventing Mental Disorders in Children: A Public Health Priority." *Canadian Journal of Public Health* 98(3): 174–78.

Weir, E. 2001. "The Health Impact of Bullying." *Canadian Medical Association Journal* 165(9): 1249.

Weir, E. 2005. "Preventing Violence in Youth." *Canadian Medical Association Journal* 172(10): 1291–92.

Williams, K., L. Rivera, R. Neighbours and V. Reznik. 2007. "Youth Violence Prevention Comes of Age: Research, Training and Future Directions." *Annual Review of Public Health* 28: 195–211.

Wilson, S.J. and M.W. Lipsey. 2007. "School-Based Interventions for Aggressive and Disruptive Behavior: Update of a Meta-Analysis." *American Journal of Preventive Medicine* 33(2 Suppl.): S130–43.

Wilson, S.J., M.W. Lipsey and J.H. Derzon. 2003. "The Effects of School-Based Intervention Programs on Aggressive Behavior: A Meta-analysis." *Journal of Consulting and Clinical Psychology* 71(1): 136–49.

Wright, J.L. 2005. "Training Healthcare Professionals in Youth Violence Prevention: Closing the Gap." *American Journal of Preventive Medicine* 29(5): 296–98.

About the Authors

Robert G. Santos, PhD, is the scientific director for the Healthy Child Manitoba Office (HCMO), in Winnipeg, Manitoba; senior policy advisor to the Healthy Child Committee of Cabinet (HCCC); and senior advisor to the deputy minister of healthy living, youth and seniors, in the government of Manitoba, where he has served since 1998. In 2006, Dr. Santos was cross-appointed as an assistant professor in the Department of Community Health Sciences, Faculty of Medicine, at the University of Manitoba, and in 2008 as a research scientist at the Manitoba Centre for Health Policy (MCHP). Dr. Santos can be reached at 204-945-8670 or by e-mail at Rob.Santos@gov.mb.ca.

Mariette J. Chartier, RN, PhD, is the senior research scientist for HCMO, HCCC, government of Manitoba, and

has served there since 2003. In 2008, Dr. Chartier was cross-appointed as an assistant professor in the Department of Community Health Sciences, Faculty of Medicine, at the University of Manitoba, and as a research scientist at MCHP. She can be reached at 204-945-2537 or by e-mail at Mariette.Chartier@gov.mb.ca.

Jeanne C. Whalen, BEd, MSc, is a research and evaluation consultant. She has implemented research projects and program evaluations in educational, hospital and community settings for over two decades, including for HCMO since 1998; and is currently a project officer for the Office of Community Medicine, First Nations and Inuit Health Branch, Health Canada. She can be reached at 613-878-7814 or by e-mail at greenrock@rogers.com.

Dan Chateau, PhD, is a research scientist at MCHP and also served from 2001 to 2010 as a consultant for the Biostatistical Consulting Unit of the Department of Community Health Sciences, Faculty of Medicine, at the University of Manitoba. He can be reached at 204-975-7767 or by e-mail at Dan_Chateau@cpe.umanitoba.ca.

Leanne Boyd, MSW, RSW, is the director of policy development, research and evaluation for HCMO, HCCC, government of Manitoba, and has served there since 1976. She was a founding member of the Manitoba Children and Youth Secretariat in 1994, which set the foundations for the Healthy Child Manitoba Strategy in 2000 and the proclamation of The Healthy Child Manitoba Act in 2007. Ms. Boyd can be reached at 204-945-5447 or by e-mail at Leanne.Boyd@gov.mb.ca.