Nurturing Kindness Naturally: A Humane Education Program's Effect on the Prosocial Behavior of First and Second Graders across China

William Ellery Samuels

The City University of New York

College of Staten Island

School of Education

Department of Educational Studies

This is an accepted manuscript of an article to be published by Elsevier in the *International Journal of Educational Research*, doi: 10.1016/j.ijer.2018.08.001.

 2018. This manuscript version is made available under the CC-BY-NC-ND 4.0 license

http://creativecommons.org/licenses/by-nc-nd/4.0/

Submitted for publication on March 31, 2018

Re-submitted on July 15, 2018

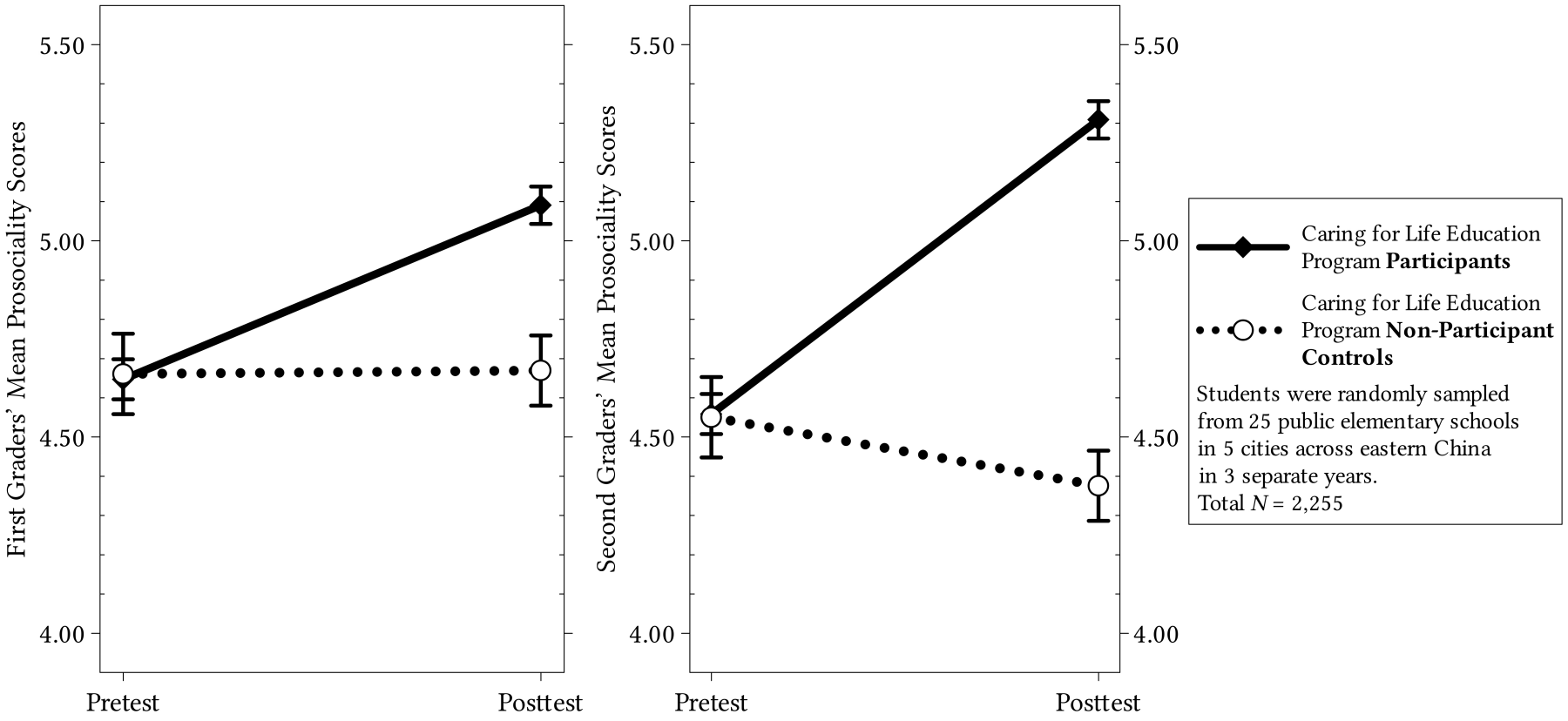
Accepted on August 5, 2018

Abstract

Research suggests that school-based programs can increase the frequencies of children's prosocial behaviors. However, extant research nearly exclusively studies relatively small studies in Western schools. Through a large evaluation conducted over 3 separate years in 25 public elementary schools in 5 cities across eastern China, we tested whether the Caring for Life humane education program—which employs animal- and nature-related content and activities—improved the prosociality of first and second grade students. Students who participated in the program displayed significantly greater gains in prosociality than similar students who didn't. Students who participated in an expanded version of the program appeared to realize even greater gains. The study supports the ability of humane education programs to benefit children in the vast but under-studied area of humane education in non-Western cultures.

Keywords: humane education, caring-for-life education, urban, elementary, prosocial behavior, first graders, intervention, People's Republic of China

Graphical Abstract



1. Introduction

Prosocial behaviors, definable as "any act that assists, benefits, or provides support for another" (Honig, 1982, p. 51), are recognized as important components of interpersonal relationships (Staub, 1971). In addition to being sought-after outcomes themselves, prosocial behaviors in children predict lower rates of internalizing and externalizing problem behaviors (Flouri & Sarmadi, 2016), lower rates of future aggression (Kokko, Tremblay, Lacourse, Nagin, & Vitaro, 2006), improved social functioning (Eisenberg, Fabes, & Spinrad, 2006; Eisenberg et al., 1996), and higher academic achievement (Caprara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000).

1.2. Development of Prosocial Behavior

The frequency of prosocial behaviors shows strong (Honig, 1982) but non-linear (House et al., 2013; Staub, 1971) growth throughout childhood. Beginning at least by the second year of life (Roth-Hanania, Davidov, & Zahn-Waxler, 2011; Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992), the frequency of these behaviors becomes increasingly affected by external events and contexts (Brownell, Svetlova, & Nichols, 2009; Flouri & Sarmadi, 2016) and is strongly guided by the quality of a child's primary relationships (Carlo, Mestre, Samper, Tur, & Armenta, 2011; Dubeau, Coutu, & Lavigueur, 2013; Newton, Laible, Carlo, Steele, & McGinley, 2014), overall socialization (Brownell, Svetlova, Anderson, Nichols, & Drummond, 2013; Carlo, 2006; Chen & French, 2008), and their interactions with others (Gross et al., 2015). Throughout childhood, their development is further guided by a growing understanding of others' emotions and goals (Thompson & Newton, 2013), the child's active involvement in collaborative experiences (Henderson, Wang, Matz, & Woodward, 2013), a growing sense of fairness and reciprocity (Sommerville, Schmidt, Yun, & Burns, 2013), and a motivation to have help given to others (Hepach, Vaish, & Tomasello, 2013).

The types of factors that affect the development of prosociality corroborate what one might expect from such a socially-oriented disposition: Prosociality largely develops through social interactions.

These interactions initially center on the quality (Newton et al., 2014; Strayer & Roberts, 2004) and extent (Dubeau, Coutu & Lavigueur, 2013) of the primary care giver's involvement with that child. However, the domain of those who affect a child's prosocial development becomes increasingly wider and more diverse as a child develops (Choi, Johnson, & Johnson, 2011; Dubeau et al., 2013; Honig, 1982, 1982; Wentzel, Barry, & Caldwell, 2004; Zahn-Waxler et al., 1992). The new relationships the child develops often add to—and don't typically supplant—the previous relationships (Flouri & Sarmadi, 2016; Gülay, 2011; Yoo, Feng, & Day, 2013), creating a growing range of interactions through which the child may grow.

1.3. Intervening to Improve Prosocial Behavior

Various programs have used interactions to encourage the development of prosociality throughout childhood. Underlining the role of parental relationships in its development, Menting, de Castro, and Matthys (2013) found strong support for benefits of a behavioral training program for new and recent parents regarding the subsequent prosocial development of their children. Others (e.g., Ogden & Hagen, 2008) have also found that parental training and support programs can promote the development of children's prosociality.

Although both home- and school-based programs can improve children's social functioning (Doescher & Sugawara, 1992), most programs are conducted in schools where children have many opportunities for social interactions and where programs can be more readily implemented and controlled than in children's homes. Bradshaw, Waasdorp, and Leaf (2012) evaluated one such widely-used program in 37 US elementary schools; the program seeks to reduce problematic behaviors in schools by addressing school culture through student support and staff training, and Bradshaw, Waasdorp, and Leaf found that it indeed reduced problem behaviors and improved prosociality. Kramer et al. (2014) and Flannery et al. (2003) also both found that school-wide programs could help most children's prosocial development. These and other programs (e.g., Caprara et al., 2014; Caprara, Kanacri,

Zuffiano, Gerbino, & Pastorelli, 2015; Fraser & Pakenham, 2008; Jordans et al., 2010; Layous, Nelson, Oberle, Schonert-Reichl, & Lyubomirsky, 2012; Menting et al., 2013; Raver et al., 2008; Samuels & Reinhartz, 2000; Schonert-Reichl et al., 2015) typically find that addressing a class's or school's social dynamics and employing a strength-based approach can promote prosocial development.

The Expanding Targets of Prosociality

As children increase the number and type of others with whom they interact, the need increases not only for more frequent prosocial behaviors, but also to direct them toward members of increasingly different outgroups (Zahn-Waxler et al., 1992). Although the role of intergroup dynamics on prosociality is not simple (Abrams, Van De Vyver, Pelletier, & Cameron, 2015; Stürmer, 2009), heterogeneous social systems may promote prosociality (Kovářík et al., 2012) and well-guided interactions with others can be used to nurture prosocial behaviors; for example, Schonert-Reichl, Smith, Zaidman-Zait, and Hertzman (2012) report that classroom visits by parents with their infants—who are outgroup members for school-aged children—increased prosocial behaviors among the students. In addition, Thielmann and Böhm (2016) found that showing prosocial behaviors to outgroup members was not done at the expense of showing it to ingroup members (or vice versa). Indeed, Thielmann and Böhm's findings suggest that improving prosociality towards one group may help children extend those same behaviors to other ones.

The presence of an infant in a classroom can be a rather salient stimulus, and this salience may have contributed to the efficacy of the program evaluated by Schonert-Reichl, Smith, Zaidman-Zait, and Hertzman (2012). Animals and animal/nature-related themes may also serve as prominent stimuli since children's attention is also often naturally piqued by animals (Belz, 2012), and people (Windhager, Atzwanger, Bookstein, & Schaefer, 2011)–especially children (Kahn, 1997; Serpell, 1999)—typically demonstrate spontaneous interest in animals. Kellert and Wilson (1993) even propose that people have an innate tendency to orient towards nature and animals.

Animal-directed empathy may encourage the development of empathy toward people since those who demonstrate greater empathy toward animals also tend to show greater empathy toward humans (Filippi et al., 2006). In addition, many animals tend to elicit empathic responses that are as strong or even stronger than the responses elicited by other people (Angantyr, Eklund, & Hansen, 2011), and attachments to animals can elicit the same neuro-endocrinological responses (Miller et al., 2009; Nagasawa, Kikusui, Onaka, & Ohta, 2009) as attachments to other people. Similarly, properly-guided interactions with animals can increase the prevalence of prosocial behaviors in children and adolescents (e.g., Martin & Farnum, 2002).

Addressing Prosociality through Humane Education

Indeed, the potential for animals to encourage outcomes like prosocial behavior has long been remarked—and even banked—upon. Unti & DeRosa (2003) note that John Locke discussed the power of teaching children moral virtues through interactions with animals. If this is not done, Locke argued, then "the custom of tormenting and killing other animals will, by degrees, harden their hearts even toward men" (1693, cited in Unti & DeRosa, 2003). Unti and DeRosa explain how this call was heeded by many in North America and Western Europe. Humane education (education that includes human-, animal-, and environment-related issues to promote care and concern) was championed along with woman suffrage and other social issues in the mid- and late-1800s, eventually becoming a core strategy of the major animal welfare organizations of the time, like the Massachusetts Society for the Prevention of Cruelty to Animals. Humane education even became compulsory in schools in several US states. The world wars highlighted the lack of hard evidence that existed on the effectiveness of humane education while the Great Depression further diverted resources and attention to other issues.

Although animal protection movements in the US experienced a revival after World War II (Unti & Rowan, 2001), humane education itself continued but with less zeal; Unti and DeRosa (2003) note

that "the difficulty of penetrating local and regional school system bureaucracies proved insurmountable for a movement with limited resources and more urgent concerns and responsibilities" (2003, p. 32). Humane education also faced increasing, countervailing social priorities uch as space-race-motivated science education and Western society's increasingly-industrialized relationship with animals and nature. For example, Jacobs and Goatly (2000) found that by the 1990s, only 2% of the activities in English language texts addressed environmental issues, of which over half did not involve any actual participation by the student.

Although social change and even missed opportunities undermined the role of humane education in the West, arguably the main impediment to its progress has been the lack of evidence about the effect it has on children's behavior and development. For better or for worse, humane education had long been expected to foster more than "just" treating animals with kindness: It has been hoped that through it children could also learn to be kind to people. Humane education has thus suffered from the lack of evidence supporting its ability to promote kindness toward people—i.e., prosociality.

However, it may be that humane education can in fact foster prosociality, and that this effect must be better studied and documented. Indeed, recently, a small but growing body of scientific evidence suggests that humane education can have this effect. Sprinkle (2008), for example, found that a school-based program that employed rescued shelter dogs could effectively increase empathy and reduce incidents of violence among elementary and middle school students. Similarly, Samuels, Meers, and Normando (2016) found that a humane education program improving upper elementary students' prosocial behavior compared with students in a control group who participated in a non-humane education program. More impressively, Piek et al. (2015) found that improvement in prosocial behavior remained both 6 and 18 months after first graders participated in an animal-centered program, compared to peers randomly assigned to the control group.

1.4. Limits of Previous Research

It is not surprising that only a few studies evaluate programs in the field where it is daunting to conduct well-controlled studies with randomized controls and objective, valid outcomes over an extended period. The studies noted above are among the few that have been able to show promise for humane education programs through well-conducted investigations in schools. However, all of these studies were conducted with North American or European children, and not all addressed children's actual behaviors. A handful of other studies did investigate the effectiveness of general prosocial programs in, e.g., Central America or Africa (Betancourt et al., 2010; Kerr, Vardhan, & Jindal, 2012), but—to our knowledge—none investigate elementary-school-based prosocial programs, let alone humane education per se.

Social and cultural differences between Western and East Asian cultures could mean there are not only differences in children's expressions of prosociality, but that there are also differences in the ways cultures respond to humane education interventions. Indeed, collectivist cultures/societies that encourage bonding to traditional family and social institutions can provide children with sources of adaptive growth (King et al., 2005; Wandersman & Nation, 1998) and protection from developmental risks (Gorman-Smith, Henry, & Tolan, 2004; Jarrett, 1997; Sheidow, Gorman-Smith, Tolan, & Henry, 2001).

Looking further within a collectivist culture, Ma (2003) reported that familial and academic factors predicted prosocial behaviors among Chinese elementary-school students, and that prosociality itself predicted good peer relationships; however, Ma found that the classroom environment did not significantly predict prosocial behaviors—a trend that does not completely replicate those found by, e.g., Choi, Johnson, and Johnson (2011) and Flouri and Sarmadi (2016) in non-collectivist cultures.

Noting the dearth of research on prosocial behaviors in non-Western societies, House et al. (2013) conducted one of the few larger-scale studies yet done with children (and adults) from a wide range

of societies, most of which were non-Western. They found that although the initial development of prosocial behaviors among young children was similar across very different cultures, subsequent development diverged considerably, likely due to social learning and the effects of assimilating various social norms.

In short, the ways in which cultures affect the development of prosocial behaviors likely varies (Chudek & Henrich, 2011) as cultures variously emphasize, e.g., the importance of collectivist norms (Kärtner, Keller, & Chaudhary, 2010; Rochat et al., 2009). The effects of prosocial programs may well vary, too.

The study presented here attempted to address both (a) the paucity of evidence of the effectiveness of prosociality programs for a non-Western child population and (b) whether including not only human- but also animal- and nature-related content and activities can serve as an effective platform. This is an ambitious set of objectives, and the current study only intends to begin such a line of research by first investigating whether such a program does have an effect and then if this effect can be attributed to the program per se and not simply, e.g., normal child development.

1.5. Overview of Current Study

We investigated the effect of the Caring for Life (CFL) program on first and second grade students across eastern China. All grades and classes in a given school either did (experimental group) or did not (control group) participate in the CFL program throughout an entire academic year. Table 1 summarizes the schools (and the number of classes in each) that participated in each of the three years.

A subset of students from each class was chosen at random to represent that class; teachers rated each of these randomly-selected students both at the beginning and the end of that academic year using the Teacher Observation of Classroom Adaptation–Checklist (TOCA-C) Prosociality subscale. The CFL program was considered effective if the students in the experimental-group schools showed

stronger growth in the year they participated in the program than did students in the control-group schools during the same year.

Considering the complexity of the data and design, this article reviews the results in three parts. The first part (Section 3) reports descriptive statistics and justifies the types of analyses used. The second part (Section 4) contains the main analyses, which investigated the effect of the CFL program as it is typically conducted. The third part (Section 5) covers ancillary analyses exploring both whether a teacher's experience with the program affected its efficacy and whether an expanded version of the program provides an additional increase in prosocial development.

2. Methods

2.1. Participants

With IRB approval from the investigator's institution and with approval from the local Chinese Ministries of Education, 2,255 students and 159 teachers from first and second grade classes in 25 public elementary schools in 5 cities across eastern China consented to participate in this study. It was made clear that participation in the CFL program itself was not contingent on participating in the study and that participating in the study was completely voluntary. Table 1 presents the numbers of participating *classes* in each year, city, and school; Table 2 presents the number of *students* who participated in each city and school each year.

Students participated for one academic year; schools that participated for more than one year had different students participate each year. None of the students in this study ever participated in the CFL program before, and they all only participated in the program only once. Students who participated as first graders, for example, did not also participate as second graders.

Schools asked to participate in the control group were chosen at random within two constraints. First, geographic and socio-economic factors were balanced between experimental and control group schools. Second, practical issues outside of the program required assigning more classes to the exper-

imental than to the control group. All schools in both the control and experimental groups agreed to participate.

Whole classes participated in the CFL program, but a subset of these participating students was chosen at random to be invited to participate in this study. We chose to study the outcomes of a subset of students in each class instead of the entire class both to standardize the number of students across classes and to maintain a representative but manageable sample of students. Pilot analyses indicated that randomly-chosen samples of 15 students from each class would yield sufficient power, so we attempted to include that many from each class in both experimental- and control-group schools. No teachers, students, or the students' parents/guardians declined to participate, although it was made clear that they were not obliged to do so.

City A is a large, metropolitan city on the north-eastern coast. City B is a small, inland city in southern China about 240 km from the coast. Cities C, D, and E are large, mostly coastal cities in the southeast.

As is common across China, participating teachers taught moral education as well as other subjects (e.g., literacy and math). Participating teachers in School X and School Y had taught the CFL program once before, in the year before they participated in this study. However, none of the students in the current study participated in any prior CFL programming. All teachers at all of the other schools (i.e., at all schools except Schools X and Y) had no prior knowledge of the CFL program. No teachers at any of the 25 participating schools indicated that they had prior knowledge of humane education before using the CFL program.

Table 1: *Number of Classes with Participating Students by Year, Group, Wave, Grade, City, and School*

				Year 1		Year 2		Year 3	
Grade	Group	City	School	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
1	Control	A	A					2	2
		B	B					10	10
			C					2	2
		C	D			2	2	4	4
			E			1			
			F			2	2		
	Experimental	A	G			1	1		
			H			2	2	3	
			I			3	3	3	
			J			2		2	
			K			2			
			L			5	4	4	
			M					2	
			N			1	1	4	
		C	O			6	6		
			P	4	4			3	3
			Q			2	2		
			R	4	4	2	2	4	4
			S	6	6				
			T	2	2				
		D	U			4	4		
			V					4	
			W					2	
		E	X	1	1	3	3	4	4

				Year 1		Year 2		Year 3	
Grade	Group	City	School	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
2	Control	A	A					2	2
		B	B					1	1
		C	D			3	3		
			F			2	4		
	Expanded Program	C	P			2	2		
	Experimental	A	H					3	
			I					4	
			J					3	
			L					5	
			M					2	
			N					2	
		C	O	1	5				
			Y	6	6				
			R	4	4	3	2		
			S			2			
			T	2	2				

Table 2

Sample Sizes and TOCA-C Prosociality Score Means and SD*s at Pre- and Posttest by Year, Group, Grade, City, and School.*

				Year 1						Year 2						Year 3					
				Pretest			Posttest			Pretest			Posttest			Pretest			Posttest		
Grade	Group	City	School	*N*	*X*	*SD*	*N*	*X*	*SD*	*N*	*X*	*SD*	*N*	*X*	*SD*	*N*	*X*	*SD*	*N*	*X*	*SD*
1	Control	A	A													32	5.07	0.93	32	4.79	0.77
		B	B													161	4.81	0.86	161	4.51	0.82
			C													30	2.85	0.31			
		C	D							30	4.57	1.04	26	5.10	0.97	67	5.22	0.73	67	4.94	0.79
			E							16	4.75	0.47									
			F							30	4.27	0.53	20	4.22	0.47						
	Grade 1 Control Total									76	4.49	0.78	46	4.72	0.90	290	4.71	1.05	290	4.66	0.82
	Experimental	A	G							16	5.50	0.35	9	5.67	0.20						
			H							31	4.54	0.64	13	4.71	0.45	39	3.88	1.01			
			I							45	4.71	1.00	45	4.84	0.46	44	4.72	0.96			
			J							30	4.31	0.64				30	4.17	0.75			
			K							32	5.12	0.56									
			L							75	5.85	0.30	48	5.73	0.58	60	4.96	0.84			
			M													30	4.75	0.71			
			N							15	6.00	0.00	8	6.00	0.00	59	4.90	1.04			
		C	O							93	4.43	0.42	93	5.28	0.65						
			P	60	4.66	0.82	60	4.88	0.69							51	4.46	1.15	51	5.25	1.04
			Q							30	4.57	1.17	30	4.73	0.93						
			R	60	4.55	1.00	60	5.46	0.50	30	3.71	1.11	30	5.01	1.14	94	4.34	0.94	94	4.93	0.62
			S	87	4.49	0.89	88	5.51	0.48												
			T	29	4.60	0.73	29	5.45	0.38												

				Year 1						Year 2						Year 3					
				Pretest			Posttest			Pretest			Posttest			Pretest			Posttest		
Grade	Group	City	School	*N*	*X*	*SD*	*N*	*X*	*SD*	*N*	*X*	*SD*	*N*	*X*	*SD*	*N*	*X*	*SD*	*N*	*X*	*SD*
1		D	U							60	4.41	1.00	57	4.58	1.05						
			V													60	5.15	0.74			
			W													30	4.20	0.94			
		E	X	10	3.84	0.86	10	5.14	0.66	45	4.88	0.55	45	4.77	1.12	70	4.09	0.56	70	4.30	0.73
	Grade 1 Experimental Total			246	4.53	0.89	247	5.32	0.60	502	4.80	0.93	378	5.07	0.91	567	4.56	0.97	567	4.79	0.88
2	Control	A	A													32	4.34	0.68	32	3.90	1.00
		B	B													17	5.08	0.80	17	4.55	0.33
		C	D							45	4.40	0.84	45	4.64	0.99						
			F							30	4.65	0.41	58	4.34	1.00						
	Grade 2 Control Total									75	4.50	0.71	103	4.47	1.00	49	4.59	0.79	49	4.12	0.88
	Expanded Program	C	P							29	3.47	0.61	23	4.80	1.01						
	Experimental	A	H													44	4.74	0.97			
			I													58	4.92	0.75			
			J													45	3.96	1.09			
			L													75	5.13	1.02			
			M													30	4.21	0.70			
			N													45	4.49	0.85			
		C	O	15	4.57	0.77	75	5.44	0.49												
			Y	30	4.49	0.84	30	5.42	0.60												
			R	60	4.48	0.93	60	5.43	0.56	44	3.81	0.99	28	4.68	1.15						
			S							30	5.18	0.61									
			T	30	4.29	0.51	30	5.43	0.65												
	Grade 2 Experimental Total			135	4.45	0.81	195	5.43	0.55	74	4.37	1.09	28	4.68	1.15	297	4.66	1.00			

2.2. Program

The Caring for Life (CFL) education program was created by ACTAsia with Nioclas Leney. ACTAsia is an international non-profit organization with presence in the United Kingdom, USA, Netherlands and Australia; ACTAsia is also registered in the People's Republic of China and actively conducts programs and outreach in several Asian, South Asian, and European countries.

The CFL program employs a variety of student-centered activities that draw on students' personal experiences to promote care for animals, people, and the environment and to help develop students' empathic self-efficacy, a likely important mediator in the expression of prosociality among children (Caprara et al., 2012). The program includes several elements found to predict subsequent prosocial behavior including cooperative learning experiences (Choi et al., 2011), simulation of animal actions (Piek et al., 2015), developing "emotional intelligence" and practicing adaptive empathic concern (Williams, O'Driscoll, & Moore, 2014), helping children develop adaptive emotional and social responses to stressors (Taylor et al., 2013), reflecting upon and celebrating previous prosocial behaviors conducted by the children (Tasimi & Young, 2016), and establishing and maintaining cooperative social norms (House et al., 2013).

Before conducting the program in their classes, teachers participate in three training sessions that familiarize them with the program's content and target outcomes as well as the pedagogical strategies used. The teachers then conduct a sequence of 12 sessions throughout an academic year (September – June), with 6 sessions per semester.

The program was conducted differently at one school, School P. At the behest of the school, the number of activities associated with each of the 12 sessions was doubled. Since the program was different at School P, we will investigate the outcomes there separately (in Section 5, Ancillary Analyses) and not include the data from School P in any other analyses.

2.3. Instrument

Participating teachers in both the experimental and control groups completed the Prosocial subscale of the Teacher Observation of Classroom Adaptation–Checklist (TOCA-C). The TOCA-C measures the frequency of developmentally adaptive and maladaptive child behaviors (Koth, Bradshaw, & Leaf, 2009). This instrument has been adapted to and extensively used in a wide range of school- and home-based studies. Versions of it are frequently used by mental health and social service practitioners to diagnose children for further treatment and intervention.

The TOCA-C asks teachers to rate how often during the previous three weeks a given student has displayed a series of behaviors, including those associated with prosociality, the domain of interest here. Specifically, it asks respondents to respond to this prompt: "*In the last three weeks*, would you say the following statements were never, rarely, sometimes, often, very often, or almost always true of this child." It then lists a series of activities, including the five items that comprise the Prosociality subscale: "Is friendly," "Is liked by classmates," "Shows empathy & compassion for others' feelings," "Is rejected by classmates," and "Harms others." Those last two items are reverse scored. Each of these items is scored on a 6-point scale where higher scores denote more frequent occurrences of a prosocial behavior (or less frequent occurrences of a non-prosocial behavior). These items scores are averaged with equal weighting, producing prosociality scores that range from 1 to 6 in increments of .2 (i.e., 1.0, 1.2, etc.).

The TOCA-C Prosocial subscale therefore measures prosociality globally. There is some indication that children may develop different types of prosocial behaviors at different rates (Dunfield, Kuhlmeier, O'Connell, & Kelley, 2011), although assessing this is beyond the scope of the current study. Instead, the outcome here is global prosociality across all occasions witnessed by the teacher.

Evidence for the valid use of the TOCA-C across several Western populations of children has been well supported by Koth, Bradshaw, and Leaf (2009). It has been used to evaluate the outcomes of

school-based interventions, including a humane education program conducted in US cities (Samuels, Meers, & Normando, 2016). The valid use of the Chinese version of the TOCA-C for assessing elementary students is supported by Samuels (2018).

2.4. Procedure

At the beginning of the academic year, exactly one week before the CFL program began at all of the experimental-group schools, participating teachers were asked to report on a randomly pre-selected subgroup of participating students' overall prosocial behavior during the past two weeks through the TOCA-C. The teachers were also asked to complete the TOCA-C for those same students at the end of the academic year, exactly one week after the completion of the entire CFL program in the experimental-group schools.

Students were assessed over one academic year only. Although the study was conducted in three consecutive years, data were collected only on students for the year in which they (or their matched, control-group peers) participated in the CFL program.

Teachers were not blind to the treatment condition: The program is intended to be conducted in their classrooms in part to facilitate generalization to other events throughout the school day and—though not measured here—to students' off-campus lives.

During the first year (year 1), all participating classes engaged in the CFL program; conditions in the field unrelated to the study required that no students be assigned to the control group during the first year. During the second and third years (years 2 and 3), classes either did or did not participate in the CFL program.

Different cohorts of students participated each year, so no student participated in the study more than once. In addition, most of the teachers who participated in the study in a given year did not participate again in any of the other two years.

Events outside of the study prevented posttest data from being collected for experimental-group schools in grade 2 during year 3. Year 3 posttest data were collected for control-group schools for both grades 1 and 2.

3. Data Exploration

Before exploring the effectiveness of the program in Section 4 (Main Analyses), we first present descriptions of the entire set of data. We also explore the best way to handle some of the complexities of these data, which include multiple academic years and various schools. Correctly handling multiple years, different schools, etc. not only yields more reliable results, it also helps establish the generalizability of the results to other contexts.

3.1. Descriptive Statistics

Table 1 presents the number of participating classes per school, city, year, and group. All of the classes at a given school of that grade participated, so Table 1 also presents how large a given school was at that grade level.

Table 2 presents the number of students about whom TOCA-C data were collected during the pre- and posttest phases for each grade, group, city, and school; this table also presents the mean (and standard deviation, *SD*) prosociality score for each of these subgroups.

The total number of students in all classes ranged from 42 to 61 (mean = 49.65, *SD* = 5.49). The mean number of students participating in this study per class was 14.06 (*SD* = 2.62). Although there was minor attrition in the number of students assessed per class, the main reason that the mean number of students per class was less than 15 was because teachers at two schools could not complete the instrument for 15 students during year 1 for reasons unrelated to this study: That year, teachers at School Y completed it for only 5 students per class and teachers at School E completed it for 10 students per class.

Table 3 presents the mean pre- and posttest prosociality scores for the first and second grade students by year and group. This table includes results both for when the program was taught normally (the Experimental groups) and for when it was taught with more activities (the Expanded Program groups).

Table 3

Pre- and Posttest TOCA-C Prosociality Scores for the Control and Experimental Group Students in Grades 1 and 2 for Each Year of the Study.

			Pretest			Posttest		
Grade	Year	Group	*N*	*X*	95% CI	*N*	*X*	95% CI
1	1	Experimental	246	4.53	0.11	247	5.32	0.07
	2	Control	74	4.49	0.18	46	4.72	0.26
		Experimental	501	4.80	0.08	366	5.07	0.09
	3	Control	268	4.71	0.13	222	4.66	0.11
		Experimental	513	4.56	0.08	162	4.79	0.14
	Grade 1 Total		1602	4.65	0.05	1043	4.98	0.05
2	1	Experimental	135	4.45	0.14	195	5.43	0.08
	2	Control	74	4.50	0.16	99	4.47	0.20
		Expanded Program	30	3.47	0.22	23	4.80	0.41
		Experimental	73	4.37	0.25	28	4.68	0.43
	3	Control	45	4.59	0.23	44	4.12	0.26
		Experimental	296	4.66	0.11			
	Grade 2 Total		653	4.51	0.07	387	4.94	0.00

3.2. Accounting for Differences in Initial Levels of Prosociality

The mean pretest scores in Table 3 suggest that students started the year with different levels of prosociality. This is not unexpected and students in different years and locations may even vary systematically in their prosociality. It is therefore quite important to consider the effects of their initial levels in the analyses and their interpretations.

We conducted a multilevel model (in which students were nested in schools, and schools were nested in cities) to test whether the grade 1 students randomly assigned to the groups in all three years of the study showed significantly varied levels of *initial* prosociality; a similar multilevel model tested for differences among the grade 2 students. The results of these analyses—presented in Table 4 for grade 1 and Table 5 for grade 2—indicated that there were some significant differences in initial prosociality scores across the three years; however, there were no significant, systematic effects of either the year or of whether a student was assigned to the experimental (including the expanded program) or control groups. The box and whisker plots presented in Figures 1 and 2 show the initial levels of prosociality among the groups and years; post hoc (Tukey HSD) analyses found that the significant differences here were between year 1 and 2 experimental groups (p = .0057) and between year 2 and 3 experimental groups (p = .00096). Therefore, although differences exist between initial prosociality scores, we believe there are no systematic differences related to these theoretically-relevant variables.

Although we found little evidence that students in different cities or schools began the study with systematically different pretest scores, it is still possible that a student's initial level of prosociality affects subsequent changes in it. For example, it may be that those who begin with high levels of prosociality have less room for subsequent growth to be measured by the TOCA-C (Stoolmiller, Eddy, & Reid, 2000), or that the program is best received by students of a certain level of initial prosociality.

Table 4

Multilevel Model Testing the Effects of Group Membership (Experimental vs. Control) and Study Year (Year 1, 2, and 3) on Grade 1 Pretest Prosociality Scores. Model uses all available grade 1 data; total N = *1602.*

Model Term	β	*df*	*t*	*p*
Group (Exp. vs. Control)	0.6126	19	1.9665	.0640
Year (Year 1, 2, or 3)	0.1863	6	1.6052	.1596
Group x Study Year	−0.4477	6	−3.5627	.0119

Table 5

Multilevel Model Testing the Effects of Group Membership (Experimental vs. Control) and Study Year (Year 1, 2, and 3) on Grade 2 Pretest Prosociality Scores. Model uses all available grade 2 data; total N = *653.*

Model Term	β	*df*	*t*	*p*
Group (Exp. vs. Control)	−0.0100	636	−0.0823	.9344
Year (Year 1, 2, or 3)	0.7714	13	1.9706	.0704
Group x Study Year	−0.4157	636	−2.2220	.0266

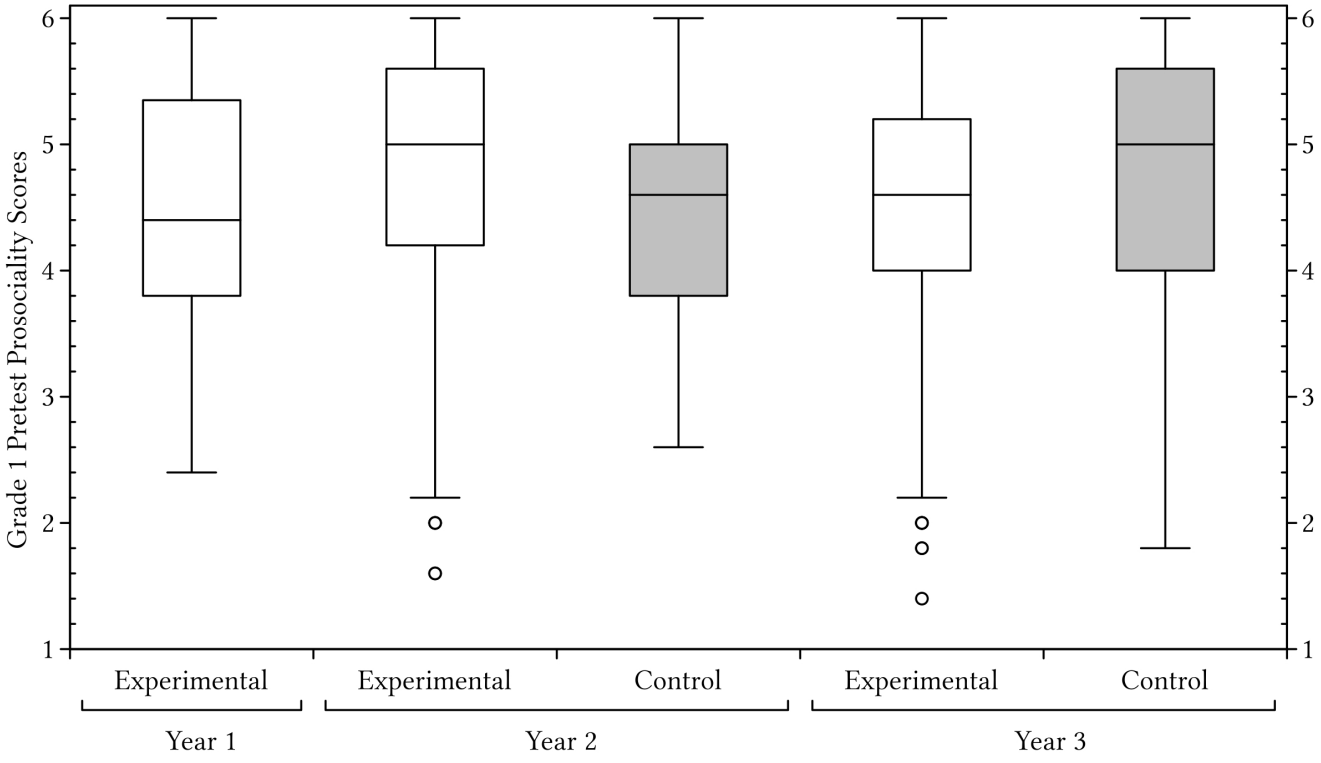


Figure 1: Box and Whisker Plots of Prosociality Pretest Scores for All Grade 1 Groups

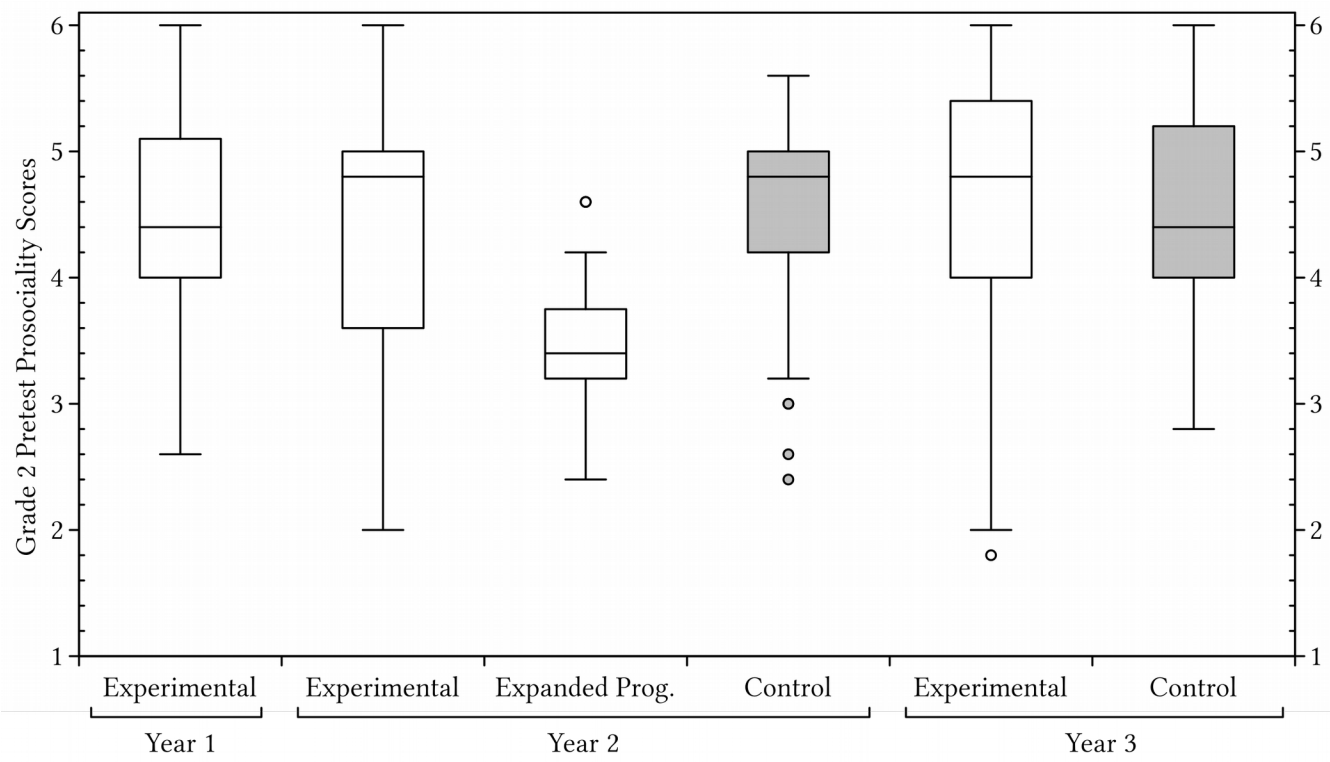


Figure 2: Box and Whisker Plots of Prosociality Pretest Scores for All Grade 2 Groups

To test the possible effect of pretest values on rates of change, we subtracted pretest prosociality scores from posttest scores for all grade 1 students and separately for all grade 2 students. We then tested the correlation between these pre-post difference scores and students' pretest scores. Pre-post differences and pretest scores indeed correlated among students both in grade 1 (r = -.578, t_{995} = -22.35, p < .001) and grade 2 (r = -.477, t_{995} = −9.32, p < .001): Students in both grades who began the year being rated as less prosocial were more likely to display stronger gains than students who began it rated as more prosocial.

Looking into this relationship further, we found that the correlation between initial level and subsequent growth was slightly larger among the experimental-group students in both grades 1 and 2 than among the control-group students in the same grade. Among the first graders, the correlation for experimental-group students was -.63 compared to -.57 for the control-group students. Among the second graders, the experimental group correlation was -.61 compared to -.57 for the control group. Overall, then, students with lower scores tended to realize greater gains, and there is some reason to believe that the CFL program itself is more effective among students with somewhat greater need.

3.3. Resultant Analytic Procedure

Given that we found significant differences among students' initial prosociality scores and given that these initial levels predict subsequent rates of gain, it is important to include as much information about students' initial and final levels of prosociality into the models as possible—and to accurately model any possible effects of year and location.

One implication is that we are ill-advised to use pre-post difference scores as our main outcome variable. Pre-post difference scores ignore most of the information in the initial pretest value, relegating it instead to error variance. Doing this would remove important information and reduce the power of our analyses. Instead, we included pretest values in all our analyses as covariates.

We also used multilevel models of change both to account for effects both within each student and between classes, schools, cities, study years and—when needed—grade levels. Specifically, time (pre- vs. posttest) was nested within student, students were nested within school, schools within city, and cities within year (grade level was also included hierarchically when analyses included both grades together). Including terms hierarchically as we do here controls better for their effects on the analyses than were we to add these variables as separate terms to, e.g., a repeated-measures ANOVA (Gelman, Hill & Yajima, 2012); this lets us focus better on the outcomes of interest with less concern for possible Type 1 or 2 errors caused by differences attributable to living in different cities, attending different schools, etc. In addition, ignoring the nested structure of variables such as these can lead to biased estimates (Chen, Kwok, Luo, & Willson, 2010). Finally, properly nesting terms is especially useful for unbalanced designs—such as when the number of students varies between classes or classes vary between schools—producing more reliable population estimates (Jeon, Lee, Hwang, & Kang, 2009).

We also used full maximum likelihood estimations in our multilevel models of change. Doing so accounts well for within-participant covariance between pre- and posttest levels in ways that can help clarify the effects of other factors in the model (Singer & Willet, 2003).

Data were not available for all participating students at both pre- and posttest. In addition, no attempt was made to collect data for control groups during year 1. The reasons for both of these were unrelated to the study and the CFL program. Even though these data are arguably missing at random (or—in some cases—missing completely at random), they can cause problems for other analyses—like repeated-measures ANOVAs—that assume homogeneous variance (Carriere, 1999). Multilevel models of change, however, handle well the heterogeneity of variance that is often caused by differently-sized groups that are exposed to different treatments; multilevel models of change easily accommodate instances where time-varying data are missing for some participants (e.g., if some students only

provide information about either the pre- or posttest). In multilevel models of change models, participants' data are used to estimate factors for which they provide data and not for those factors for which they don't (Singer & Willet, 2003).

In summary, then, appropriately nesting terms and using multilevel models of change allow for accurate, efficient, and asymptotically-unbiased use of all of the data from various years, cities, schools, students, etc. Note, however, that the power of any analysis is still limited by the size of the smallest sample (Usami, 2014), so Type 2 errors (false negatives) in particular continue to warrant attention; this matters most here for the ancillary analyses (Section 5).

Finally, the tables and figures here present raw prosociality scores to facilitate interpretation across studies and with others' work. To simplify analyses and their interpretations, however, prosociality scores were standardized within each grade before being added to the various models—or were instead standardized across both grades when both grades were included in a given model. All analyses were conducted with R, version 3.2.3 (R Core Team, 2013) interfaced through Rstudio version 1.0.153. R packages used included psych (Revelle, 2014), nlme (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2015) and multcomp (Hothorn et al., 2017).

4. Main Analyses

The main goal of this paper is to assess whether participating in the CFL program is associated with increased development of prosociality in first and second grade students in China. The primary measure of this outcome is whether students in the experimental group, who participated in the program, were reported by their teachers to have larger pre-post gains in prosociality than students in the control group. This difference manifests as a test of the interaction between group (experimental vs. control) and time (pre- vs. posttest).

Multilevel models provide more powerful estimates when all available data are included in them. Specifically, classes of students in years 2 and 3 were randomly assigned within constraints to either the experimental or control group; all students in year 1, however, were assigned to the experimental group. We will therefore test two models: one that includes data from all three years (yielding the most powerful estimates) and one that includes data from only years 2 and 3 (using only the most direct comparisons). Tests of the CFL program in these main analyses do not include those students who participated in the expanded program; they are addressed in Section 5.2.

4.1. Analyses of Full Data Set

4.1.1. Grade 1.

Figure 3 presents the results among first grade participants for each year (with the thin, grey lines) as well as for all three years combined (with the heavy, black lines). This figure depicts the variation between years for each of the groups, but it also shows that experimental-group students realized stronger growths in prosociality than control-group students.

Table 6 presents the results of the model testing the relationships shown in Figure 3. This table indicates that there is a significant group x time interaction effect (t_{993} = 3.80, p = .0002). This interaction confirms that the growth of prosociality over the given year was significantly larger among experimental-group students than among control-group students. In other words, we found strong evidence that prosociality developed more rapidly among grade 1 students who participated in the CFL program than among similar students who didn't.

(All of the tables presenting analyses—including Table 6—also provide the results for the tests of the main effects of group and time. These two main effect terms do not address the goals of this investigation; they were included in the models to isolate their effects from the term of interest (the group x time interaction term); they are reported here simply for completeness.)

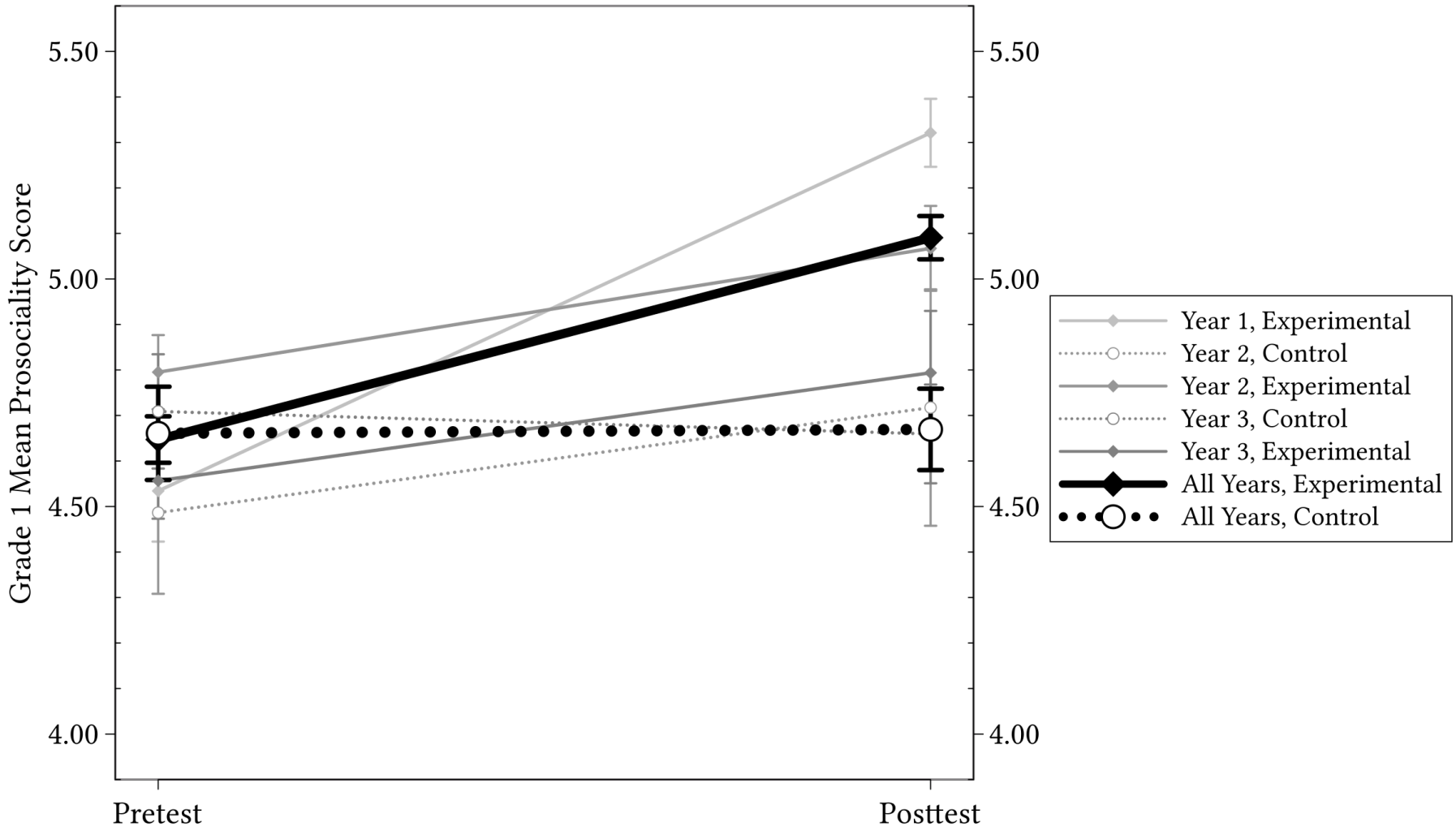


Figure 3: Study 2, Grade 1 Pre-Post TOCA-C Prosocial Behavior Scores for Students in the Experimental and Control Groups

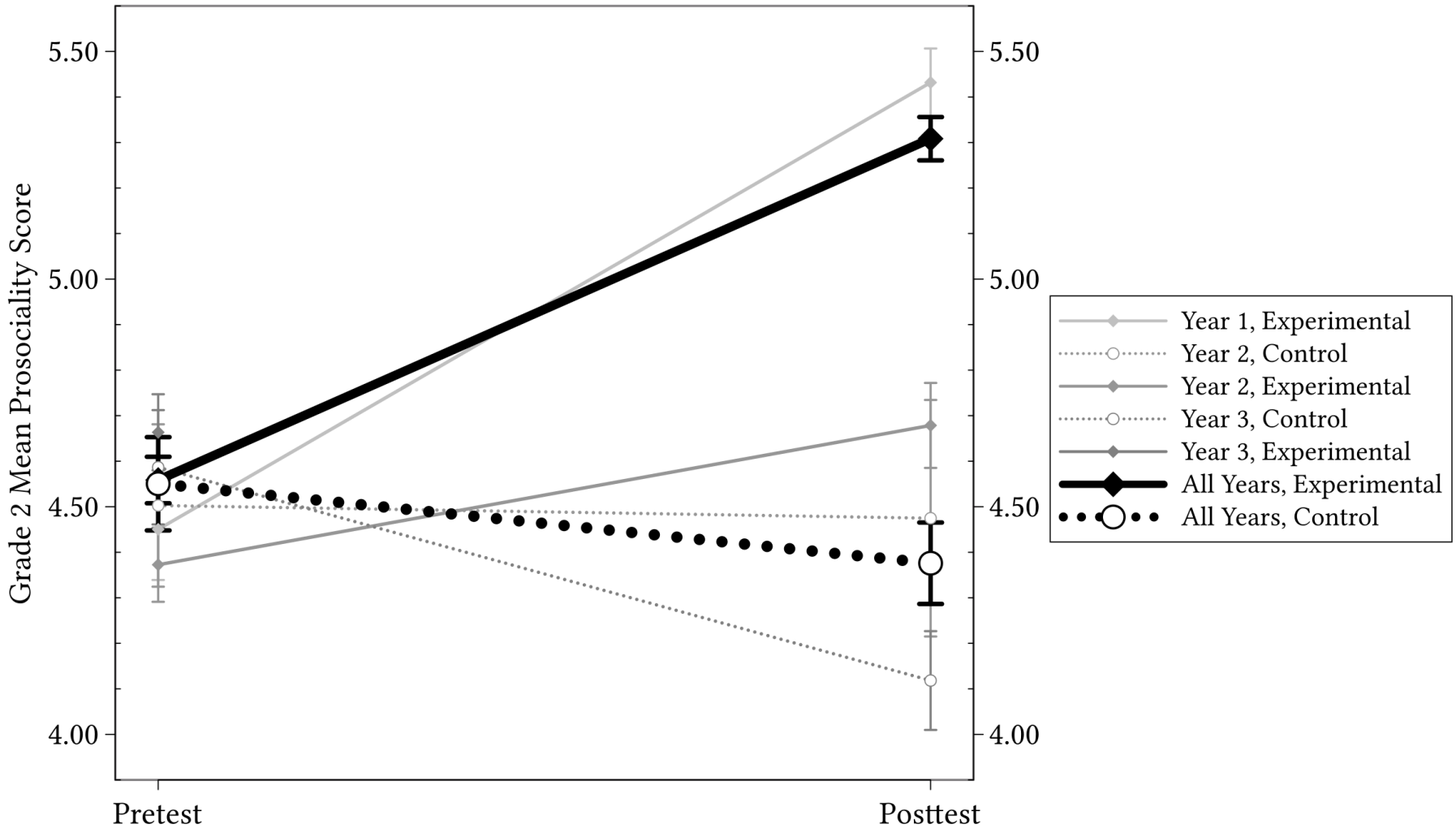


Figure 4: Study 2, Grade 2 Pre-Post TOCA-C Prosocial Behavior Scores for Students in the Experimental and Control Groups

4.1.2. Grade 2.

We also found a significant group x time interaction term among the grade 2 students (t_{295} = 7.78, p < .0001), indicating that second graders who participated in the program showed stronger growth in the prosociality than those who did not. Figure 4 and Table 7 summarize these results.

Table 6

Multilevel Model of Change Testing the Effects of Group (Experimental vs. Control) and Time (Pre- vs. Posttest) on Grade 1 Prosociality Scores. Model uses data from all study years (i.e., years 1, 2, and 3); total N *= 2644 (includes waves nested in student).*

Model Term	β	*df*	*t*	*p*
Group (Exp. vs. Control)	−0.2180	24	−1.5395	.1368
Time (Pre- vs. Posttest)	−0.1758	993	−1.0859	.2778
Group x Time	0.7150	993	3.8041	.0002

Table 7

Multilevel Model of Change Testing the Effects of Group (Experimental vs. Control) and Time (Pre- vs. Posttest) on Grade 2 Prosociality Scores. Model uses data from all study years (i.e., years 1, 2, and 3); total N *= 1044 (includes waves nested in student).*

Model Term	β	*df*	*t*	*p*
Group (Exp. vs. Control)	−0.2406	13	−1.7419	.1051
Time (Pre- vs. Posttest)	−0.2081	295	−1.7960	.0735
Group x Time	1.1242	295	7.7833	< .0001

4.2. Analyses Excluding Year 1

In years 2 and 3, students were randomly assigned to either the experimental or the group control. However, during year 1, students were only assigned to the experimental group. To ensure that including data from year 1 did not adversely affect the results, we repeated the analyses without them.

4.2.1. Grade 1.

Excluding year 1 students did not change the pattern of the results. As shown in Table 8, the group x time interaction (t_{748} = 3.56, p = .0004) was significant, indicating that the program was effective.

4.2.2. Grade 2.

Once again, there was a significant interaction between group and time with second grade students when year 1 participants were excluded (t_{135} = 2.71, p = .0077, Table 9).

Table 8

Multilevel Model of Change Testing the Effects of Group (Experimental vs. Control) and Time (Pre- vs. Posttest) on Grade 1 Prosociality Scores. Model uses data from study years 2 and 3 (i.e., year 1 is excluded); total N *= 2152 (includes waves nested in student).*

Model Term	β	*df*	*t*	*p*
Group (Exp. vs. Control)	−0.1618	21	−1.0083	.3248
Time (Pre- vs. Posttest)	−0.2350	748	−1.6043	.1091
Group x Time	0.6182	748	3.5617	.0004

Table 9

Multilevel Model of Change Testing the Effects of Group (Experimental vs. Control) and Time (Pre- vs. Posttest) on Grade 2 Prosociality Scores. Model uses data from study years 1 and 2 (i.e., year 3 is excluded); total N *= 659 (includes waves nested in student).*

Model Term	β	*df*	*t*	*p*
Group (Exp. vs. Control)	−0.1354	9	−0.8200	.4334
Time (Pre- vs. Posttest)	−0.2692	135	−1.7954	.0748
Group x Time	0.7004	135	2.7076	.0077

4.3. Discussion of Main Analyses

Participating in the CFL program resulted in reliable increases in the frequency of prosocial behaviors among both first and second grade students. These results were robust against differences in initial levels of prosociality and against any effects of the location (city or school) or the year in which the program was conducted.

5. Ancillary Analyses

Since the CFL program was conducted at many schools and at different times, it may not be surprising that there was some variation in how it was conducted. In part, this lends support to the generalizability of the outcomes. In part, it allows—even requires—us to investigate the possible effects of systematic differences in the way the program was conducted. Therefore, in Section 5 we investigated the possible effects of two variations to the typical procedure: The teachers' levels of experience with the CFL program and the amount of CFL programming conducted.

5.1. Teachers' Experience with the CFL Program

In all schools but two, the teachers had never conducted the CFL program before. Teachers participated in three training sessions to familiarize them to the program before they taught it, but otherwise they all reported that they had no special knowledge of animal- or nature-related issues outside

of their normal teacher preparation and lesson planning. Teachers at two schools—School X and Y—had previously conducted the CFL program once with other students (none of those students are included in the current study). The CFL program is designed to provide teachers with the content knowledge needed to effectively conduct the program; nonetheless, it is quite possible that this previous experience with the program affected their ability to promote prosociality.

We tested this by comparing the prosociality of students who were taught by "experienced" teachers (those in Schools X and Y who had taught the CFL program the previous year) with those who were taught by "inexperienced" teachers (who had not taught it before). We included the data from all three years, and we analyzed both grades together and then only the first grade students; only 30 second-grade students were taught by "experienced" teachers—and all of these were taught during the first year of the study—so we could not conduct reliable tests on just the second graders. We also did not include students who participated in the expanded version of the CFL program discussed later in this article.

5.1.1. Analysis of Both Grades

When both grades were combined, 143 students had been taught by "experienced" teachers: 10 first graders and 30 second graders in year 1, 45 first graders in year 2, and 60 first graders in year 3.

We did not find a significant experience x time interaction ($\beta = -0.021$, $t_{909} = -0.100$, $p = .920$). Having taught the CFL program once before did not appreciably improve how effectively teachers taught the program, compared to teachers who had never taught it before.

5.1.2. Analyses of Grade 1 Alone

Similarly, we found no significant effect of teachers' levels of experience when we looked at only the first grade students. I.e., the experience x time interaction was not significant ($\beta = -0.045$, $t_{747} = -0.16$, $p = .875$).

5.2. Amount of CFL Programming

During year 2, all of the second grade classes in one school, School P, participated in an "expanded" version of the CFL programs' standard curriculum. Students at this school participated in twice the number of program-based activities. The CFL program includes suggested, additional activities that educators could use to expand upon given topics. The administration of School P worked with a representative of ACTAsia to provide all of these additional activities to all participating students at their school. Students participated in these additional activities when the corresponding unit was being covered. Therefore, the content covered at School P remained largely the same, but it was covered in more detail with twice as much time spent engaging students in activities. No other school included any additional activities; at all other schools effort was made to conduct the program uniformly.

Including additional student activities at School P's request allowed for a preliminary assessment of whether expanding students' involvement with the program could lead to stronger increases in prosociality compared to those increases realized through the standard program.

5.2.1. Analysis of the Full Data Set

We first tested whether students who participated in the expanded program condition differed in their initial levels of prosociality from students assigned to the program's standard curriculum. When all three years were included (Table 10), there was indeed a significant difference in the pretest prosociality scores between the students who participated in the expanded program and those who participated in the standard program ($t_{13} = -2.99$, $p = .010$). Looking at Figure 5, we see that those who participated in the expanded program began with significantly lower levels of prosociality than those who participated in the standard program.

We must therefore be cautious in interpreting the results here. Not only was the expanded curriculum conducted at only one, non-randomly chosen school, but—in general—students who start

with initially lower levels of prosociality tend to subsequently show greater gains (as shown by the significant correlations between pretest scores and pre-post difference scores in Section 3.2).

Keeping these provisos in mind, we found a significant group x time interaction (t_{185} = 2.34, p = .020). Second graders who participated in the expanded program displayed significantly stronger growth in prosociality than did second graders who participated in the standard program.

Table 10

Multilevel Model of Change Testing the Effects of Participating in the Standard Versus Expanded Program (Group) and Time (Pre- vs. Posttest) on Grade 2 Prosociality Scores. Model uses data from all study years (i.e., years 1, 2, and 3); total N *= 782 (includes waves nested in student).*

Model Term	β	*df*	*t*	*p*
Group (Standard vs. Expanded Program)	−1.2316	10	−2.7884	.0192
Time (Pre- vs. Posttest)	0.9624	185	18.1468	.0000
Group x Time	0.3422	185	2.3472	.0200

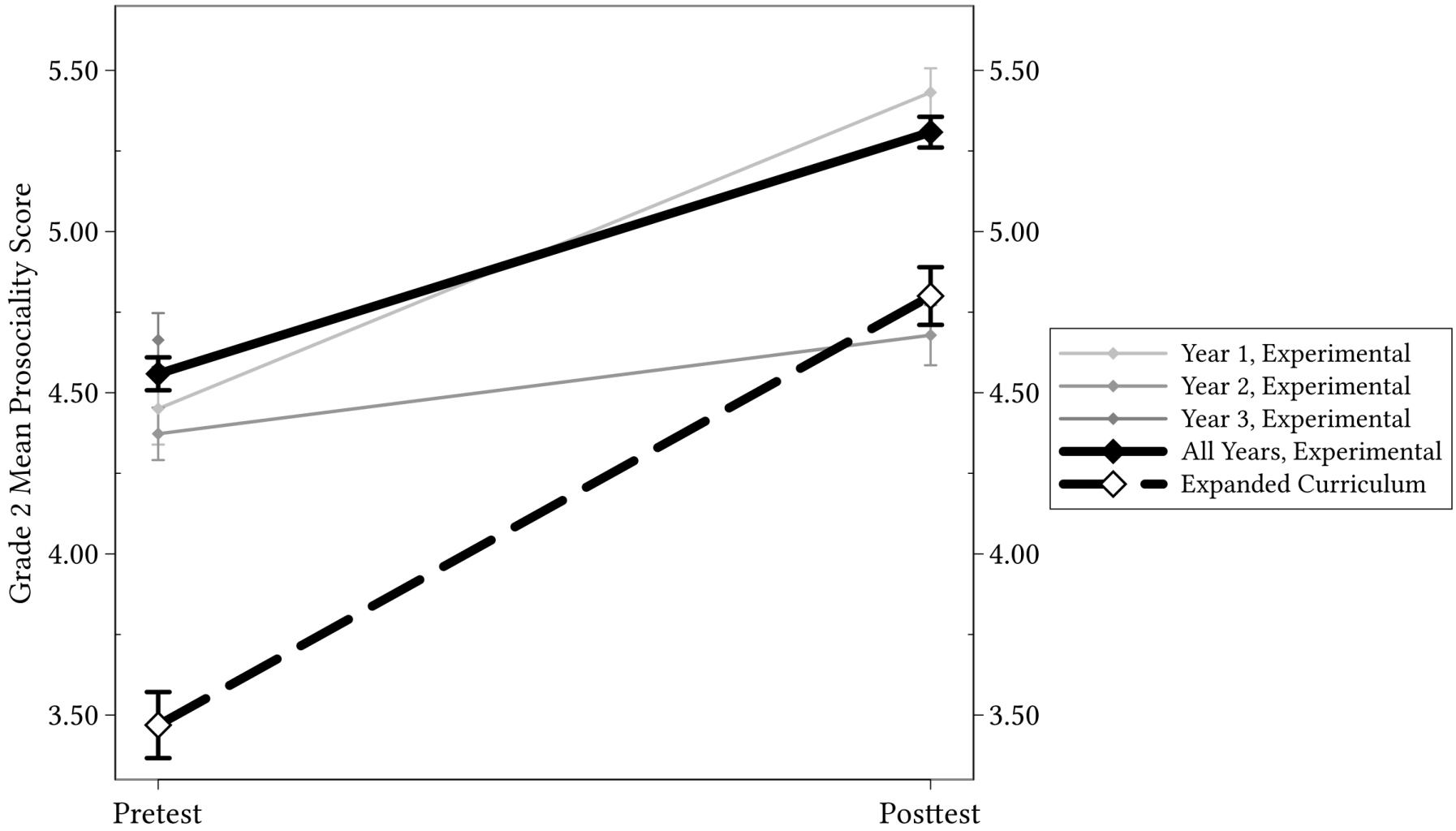


Figure 5: Study 3, Grade 2 Pre-Post TOCA-C Prosocial Behavior Scores for Students in the Experimental and Expanded Program Groups

5.2.2. Analyses of Year 2 Alone

Including all three years provides more powerful and reliable analyses of the expanded program. Nonetheless, one could argue that it is more valid to compare the expanded program with only those who participated in the standard program during that same year.

We therefore compared the growth of prosociality of those who participated in the expanded curriculum against those who participated in the standard program the same year (i.e., year 2). The benefit of the expanded curriculum still held, as indicated by the still-significant group x time interaction term (t_{50} = 2.01, p = .0498).

Those who participated in the expanded program in year 2 also did not begin at different, initial levels of prosociality from year 2 students who participated in the standard program (t_4 = −2.53, p = .064). We therefore found that participating in the expanded program produced stronger gains even when they are compared against students from the same year who also started with more similar levels of prosociality.

Table 11

Multilevel Model of Change Testing the Effects of Participating in the Standard Versus Expanded Program (Group) and Time (Pre- vs. Posttest) on Grade 2 Prosociality Scores. Model uses data only from year 2 (i.e., years 1 and 3 are excluded); total N *= 156 (includes waves nested in student).*

Model Term	β	*df*	*t*	*p*
Group (Standard vs. Expanded Program)	−1.2316	2	−2.0213	.1806
Time (Pre- vs. Posttest)	0.8598	50	5.4690	.0000
Group x Time	0.4566	50	2.0105	.0498

5.3. Discussion of Ancillary Analyses

We did not find evidence that students taught by "experienced" teachers (who had taught the CFL program once before) developed prosociality differently than students taught by "inexperienced" teachers (who had not taught the CFL program before). The program can be effectively taught by teachers who have not taught it before and who have limited experience with Caring for Life issues.

We did, however, find evidence that increasing the exposure to the program led to a larger gain in prosociality. Students who participated in an expanded version of the CFL program that had twice the amount of student-centered activities showed significantly steeper increases than students who participated in the standard program. The expanded program, however, was only conducted with second graders at one school that was not chosen at random.

Taken as a whole, the results of these ancillary analyses encourage further investigation into the effect of the length of humane education programs or the amount of student-centered activities they include. The role of a teacher's experience may be less critical for well-structured programs, but—of course—also still warrants further attention.

6. Overall Discussion

Students who participated in the CFL program developed prosociality more strongly than students in closely-matched control groups who did not participate in any additional programming. Participating in the program had a strong effect, especially among second grade students and especially among those who began the study with relatively lower initial levels of prosociality.

We also found evidence that second grade students who experienced additional learning opportunities through an expanded program showed additional gains. We did not find evidence that having taught the program once before helped teachers realize stronger gains among their students.

Together, these results indicate that a supplemental, school-based program that includes animal- and nature-related content and activities can increase lower elementary students' prosociality. These

results were found in many schools across eastern China. The analyses we used protect our conclusions well against potential confounds like student's school or city, the year in which it was studied, or even initial levels prosociality.

It is worth noting that we cannot say *what* aspects of the program were most important to elicit its hearty effect. For example, we were not able to test the extent to which animal- and nature-related content per se led to this increase over, say, the student-centered way the program was taught.

We *can* say that these results support a small but growing body of research that finds that including animals and nature in educational programs promotes prosociality, empathy, and general concern for others. Like the programs that both Piek et al. (2015) and Samuels, Meers, and Normando (2016) found effective, the currently-evaluated program included themes and activities that involved animals and nature. Those programs and this CFL program all included activities that were intended to help the animals (and people) with whom the children interacted or witnessed in their own lives, but none of the programs included direct interactions with animals. Therefore, addressing animal- and nature-related content per se appears sufficient to promote prosociality. Including animals directly may be additionally effective (Sprinkle, 2008), but doing so comes with additional safety and class management concerns.

6.1. Conclusions

The current study adds to the body of research supporting the effects of humane education programs on children's prosocial behaviors. It also extends the evidence to urban students in the Far East. At its most basic, this acknowledges the importance of considering a global audience when addressing global issues.

More directly relevant to prosocial development, the study addresses whether similar results are possible among non-Western, more collectivist cultures where pedagogical styles can differ from

Western styles (Li, Rao, & Tse, 2012; Shih, 1999). Similar results are indeed possible. A well-designed humane education program can realize strong gains in the field—both in the West and in the East.

6.2. Limitations

Feasibility necessitated that the teachers who conducted the program also evaluated the students. It seems unlikely that all of the teachers would be affected by response bias or that any biases would sway results in all the same direction: Although there may be an inclination to respond positively about the program, it's unlikely that teachers remembered the scores they gave to students over 10 months ago and the key outcome here was growth over time. In addition, we have no reason to believe that the teachers in the control group were any less inclined to rate the development of the students under their care differently. Nonetheless, we cannot discount the effect of response biases nor indicate what systematically positive or negative effect they may have had.

Students in the controls group participated in no additional programming. Although others (e.g., Samuels et al., 2016) found similar effects for another humane education program when control-group students also participated in additional programming (that was unrelated to prosociality), we cannot discern how much of the growth here was due to simple involvement in a program.

The cities in which the students lived ranged across eastern China. The results of this study, then, generalize best to other urban-dwelling children. It may well be that children living in less urban areas would respond differently.

Acknowledgements

We would like to thank the participating students, their families, and their participating teachers for agreeing to take part in this study. We would also like to thank both the schools' administrators and ACTAsia for their permissions to conduct this study. Finally, we want to heartily thank Kerenza Vlastou, Isobel Zhang, and Su Meiru for their priceless contributions to this logistically-challenging study.

References

Abrams, D., Van De Vyver, J., Pelletier, J., & Cameron, L. (2015). Children's prosocial behavioural intentions towards outgroup members. *British Journal of Developmental Psychology*, *33*(3), 277–294. https://doi.org/10.1111/bjdp.12085

Angantyr, M., Eklund, J., & Hansen, E. M. (2011). A comparison of empathy for humans and empathy for animals.(Report). *Anthrozoös*, *24*(4), 369–378. https://doi.org/10.2752/175303711X13159027359764

Belz, P. (2012). Exploring plants, insects, and animals: Opportunities for cultivating empathy in children. *Exchange: The Early Childhood Leaders' Magazine Since 1978*, (205), 67–69.

Betancourt, T. S., Borisova, I. I., Williams, T. P., Brennan, R. T., Whitfield, T. H., de la Soudiere, M., … Gilman, S. E. (2010). Sierra Leone's former child soldiers: A follow-up study of psychosocial adjustment and community reintegration. *Child Development*, *81*(4), 1077–1095.

Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2012). Effects of School-Wide Positive Behavioral Interventions and Supports on child behavior problems. *Pediatrics*, *130*(5), e1136–e1145. https://doi.org/10.1542/peds.2012-0243

Brownell, C. A., Svetlova, M., Anderson, R., Nichols, S. R., & Drummond, J. (2013). Socialization of early prosocial behavior: Parents' talk about emotions is associated with sharing and helping in toddlers. *Infancy*, *18*(1), 91–119. https://doi.org/10.1111/j.1532-7078.2012.00125.x

Brownell, C. A., Svetlova, M., & Nichols, S. (2009). To share or not to share: When do toddlers respond to another's needs? *Infancy*, *14*(1), 117–130. https://doi.org/10.1080/15250000802569868

Caprara, G. V., Alessandri, G., & Eisenberg, N. (2012). Prosociality: The contribution of traits, values, and self-efficacy beliefs. *Journal of Personality and Social Psychology*, *102*(6), 1289–1303. https://doi.org/10.1037/a0025626

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. (2000). Prosocial foundations of children's academic achievement. *Psychological Science*, *11*, 302–306. https://doi.org/10.1111/1467-9280.00260

Caprara, G. V., Kanacri, B. P. L., Gerbino, M., Zuffianò, A., Alessandri, G., Vecchio, G., … Bridglall, B. (2014). Positive effects of promoting prosocial behavior in early adolescence: Evidence from a school-based intervention. *International Journal of Behavioral Development*, *38*(4), 386–396. https://doi.org/10.1177/0165025414531464

Caprara, G. V., Kanacri, B. P. L., Zuffiano, A., Gerbino, M., & Pastorelli, C. (2015). Why and how to promote adolescents' prosocial behaviors: Direct, mediated and moderated effects of the CEPIDEA school-based program. *Journal of Youth and Adolescence*, *44*(12), 2211–2229. https://doi.org/10.1007/s10964-015-0293-1

Carlo, G. (2006). Care-based and altruistically based morality. In M. Killen & J. Smetana (Eds.) *Handbook of Moral Development* (pp. 551–579). Mahwah, NJ: Erlbaum.

Carlo, G., Mestre, M. V., Samper, P., Tur, A., & Armenta, B. E. (2011). The longitudinal relations among dimensions of parenting styles, sympathy, prosocial moral reasoning, and prosocial behaviors. *International Journal of Behavioral Development*, *35*(2), 116–124.

Carriere, K. C. (1999). Methods for repeated measures data analysis with missing values. *Journal of Statistical Planning and Inference*, *77*(2), 221–236.

Chen, Q., Kwok, O.-M., Luo, W., & Willson, V. L. (2010). The impact of ignoring a level of nesting structure in multilevel growth mixture models: A Monte Carlo study. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(4), 570–589. https://doi.org/10.1080/10705511.2010.510046

Chen, X., & French, D. C. (2008). Children's social competence in cultural context. *Annual Review of Psychology*, *59*, 591–616. https://doi.org/10.1146/annurev.psych.59.103006.093606

Choi, J., Johnson, D. W., & Johnson, R. (2011). Relationships among cooperative learning experiences, social interdependence, children's aggression, victimization, and prosocial behaviors. *Journal of Applied Social Psychology*, *41*(4), 976–1003. https://doi.org/10.1111/j.1559-1816.2011.00744.x

Chudek, M., & Henrich, J. (2011). Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*, *15*(5), 218–226.

Doescher, S. M., & Sugawara, A. I. (1992). Impact of prosocial home- and school-based interventions on preschool children's cooperative behavior. *Family Relations*, *41*(2), 200–204. https://doi.org/10.2307/584833

Dubeau, D., Coutu, S., & Lavigueur, S. (2013). Links between different measures of mother/father involvement and child social adjustment. *Early Child Development and Care*, *183*(6), 791–809. https://doi.org/10.1080/03004430.2012.723442

Dunfield, K., Kuhlmeier, V. A., O'Connell, L., & Kelley, E. (2011). Examining the diversity of prosocial behavior: Helping, sharing, and comforting in infancy. *Infancy*, *16*(3), 227–247. https://doi.org/10.1111/j.1532-7078.2010.00041.x

Eisenberg, N., Fabes, R. A., Karbon, M., Murphy, B. C., Wosinski, M., Polazzi, L., … Juhnke, C. (1996). The relations of children's dispositional prosocial behavior to emotionality, regulation, and social functioning. *Child Development*, *67*(3), 974–992. https://doi.org/10.1111/j.1467-8624.1996.tb01777.x

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Chapter 11: Prosocial Development. In W. Damon, R. M. Lerner, & N. Eisenberg (Eds.) *Handbook of child psychology: Vol. 3. Social, emotional, and personality development (6th Ed.)* (pp. 646–718). Wiley.

Filippi, M., Riccitelli, G., Falini, A., Salle, F. D., Vuilleumier, P., Comi, G., & Rocca, M. A. (2006). The brain functional networks associated to human and animal suffering differ among omnivores, vegetarians and vegans. *PLOS One*. https://doi.org/10.1371/journal.pone.0010847

Flannery, D. J., Vazsonyi, A. T., Liau, A. K., Guo, S., Powell, K. E., Atha, H., … Embry, D. (2003). Initial behavior outcomes for the PeaceBuilders universal school-based violence prevention program. *Developmental Psychology*, *39*(2), 292–308. https://doi.org/10.1037/0012-1649.39.2.292

Flouri, E., & Sarmadi, Z. (2016). Prosocial behavior and childhood trajectories of internalizing and externalizing problems: The role of neighborhood and school contexts. *Developmental Psychology*, *52*(2), 253–258. https://doi.org/10.1037/devo000076

Fraser, E., & Pakenham, K. (2008). Evaluation of a resilience-based intervention for children of parents with mental illness. *Australian & New Zealand Journal of Psychiatry*, *42*(12), 1041–1050.

Gelman, A., Hill, J. & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, *5*, 189-211. doi: 10.1080/19345747.2011.618213

Gorman-Smith, D., Henry, D. B., & Tolan, P. H. (2004). Exposure to community violence and violence perpetration: The protective effects of family functioning. *Journal of Clinical Child and Adolescent Psychology*, *33*(3), 439–449.

Gorman-Smith, D., Tolan, P. H., Henry, D. B., & Florsheim, P. (2000). Patterns of family functioning and adolescent outcomes among urban African American and Mexican American families. *Journal of Family Psychology*, *14*(3), 436–457.

Gross, R. L., Drummond, J., Satlof-Bedrick, E., Waugh, W. E., Svetlova, M., & Brownell, C. A. (2015). Individual differences in toddlers' social understanding and prosocial behavior: Disposition or socialization? *Frontiers in Psychology*, *6*, 600–610. https://doi.org/10.3389/fpsyg.2015.00600

Gülay, H. (2011). Assessment of the prosocial behaviors of young children with regard to social development, social skills, parental acceptance-rejection and peer relationships. *Journal of Instructional Psychology*, *38*(3/4), 164–172.

Henderson, A. M. E., Wang, Y., Matz, L. E., & Woodward, A. L. (2013). Active experience shapes 10-month-old infants' understanding of collaborative goals. *Infancy*, *18*(1), 10–39. https://doi.org/10.1111/j.1532-7078.2012.00126.x.

Hepach, R., Vaish, A., & Tomasello, M. (2013). A new look at children's prosocial motivation. *Infancy*, *18*(1), 67–90. https://doi.org/10.1111/j.1532-7078.2012.00130.x.

Honig, A. S. (1982). Prosocial development in children. *Young Children*, *37*(5), 51–62.

Hothorn, T., Bretz, F., Westfall, P., Heiberger, R. M., Schuetzenmeister, A., & Scheibe, S. (2017). *multcomp: Simultaneous Inference in General Parametric Models*. Version 1.4-8. Retrieved from http://multcomp.r-forge.r-project.org/

House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., … Laurence, S. (2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences*, *110*(36), 14 586–14 591. https://doi.org/10.1073/pnas.1221 217 110

Jacobs, G. M., & Goatly, A. (2000). The treatment of ecological issues in ELT coursebooks. *ELT Journal*, *54*(3), 256-264.

Jarrett, R. L. (1997). African American family and parenting strategies in impoverished neighborhoods. *Qualitative Sociology*, *20*(2), 275–288.

Jeon, M.-J., Lee, G., Hwang, J.-W., & Kang, S.-J. (2009). Estimating reliability of school-level scores using multilevel and generalizability theory models. *Asia Pacific Education Review*, *10*(2), 149–158. https://doi.org/10.1007/s12 564-009-9014-3

Jordans, M. J. D., Komproe, I. H., Tol, W. A., Kohrt, B. A., Luitel, N. P., Macy, R. D., & de Jong, J. T. V. M. (2010). Evaluation of a classroom-based psychosocial intervention in conflict-affected Nepal: A cluster randomized controlled trial. *Journal of Child Psychology and Psychiatry*, *51*(7), 818–826. https://doi.org/10.1111/j.1469-7610.2010.02209.x

Kahn, P. H. (1997). Developmental psychology and the biophilia hypothesis: Children's affiliation with nature. *Developmental Review*, *17*(1), 1–61. https://doi.org/10.1006/drev.1996.0430

Kärtner, J., Keller, H., & Chaudhary, N. (2010). Cognitive and social influences on early prosocial behavior in two sociocultural contexts. *Developmental Psychology*, *46*(4), 905–914. https://doi.org/10.1037/a0019718

Kellert, S. R., & Wilson, E. O. (Eds.). (1993). *The Biophilia hypothesis*. Washington, D.C.: Washington, D.C. : Island Press.

Kerr, J., Vardhan, M., & Jindal, R. (2012). Prosocial behavior and incentives: Evidence from field experiments in rural Mexico and Tanzania. *Ecological Economics*, *73*, 220–227. https://doi.org/10.1016/j.ecolecon.2011.10.031

King, G., McDougall, J., DeWit, D., Hong, S., Miller, L., Offord, D., … LaPorta, J. (2005). Pathways to children's academic performance and prosocial behaviour: Roles of physical health status, environmental, family, and child factors. *International Journal of Disability, Development & Education*, *52*(4), 313–344. https://doi.org/10.1080/10349120500348680

Kokko, K., Tremblay, R. E., Lacourse, E., Nagin, D. S., & Vitaro, F. (2006). Trajectories of prosocial behavior and physical aggression in middle childhood: Links to adolescent school dropout and physical violence. *Journal of Research on Adolescence*, *16*, 403–428. https://doi.org/10.1111/j.1532-7795.2006.00500.x

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2009). Teacher Observation of Classroom Adaptation–Checklist: Development and factor structure. *Measurement and Evaluation in Counseling and Development*, *42*(1), 15–30. https://doi.org/10.1177/0748175609333560

Kovářík, J., Brañas-Garza, P., Cobo-Reyes, R., Espinosa, M. P., Jiménez, N., & Ponti, G. (2012). Prosocial norms and degree heterogeneity in social networks. *Physica A: Statistical Mechanics and Its Applications*, *391*(3), 849–853. https://doi.org/10.1016/j.physa.2011.08.040

Kramer, T. J., Caldarella, P., Young, K. R., Fischer, L., & Warren, J. S. (2014). Implementing "Strong Kids" school-wide to reduce internalizing behaviors and increase prosocial behaviors. *Education and Treatment of Children*, *37*(4), 659–680.

Layous, K., Nelson, S. K., Oberle, E., Schonert-Reichl, K. A., & Lyubomirsky, S. (2012). Kindness counts: Prompting prosocial behavior in preadolescents boosts peer acceptance and well-being. *PLoS ONE*, *7*(12), e51 380. https://doi.org/10.1371/journal.pone.0051380

Li, H., Rao, N., & Tse, S. K. (2012). Adapting Western Pedagogies for Chinese Literacy Instruction: Case Studies of Hong Kong, Shenzhen, and Singapore Preschools. *Early Education and Development*, *23*(4), 603–621. https://doi.org/10.1080/10 409 289.2010.536441

Ma, H. K. (2003). The relation of academic achievement, family and classroom social environment, and peer interactions to prosocial and antisocial behavior of Chinese children. *Psychologia*, *46*(3), 163–173.

Ma, L. K., Tunney, R. J., & Ferguson, E. (2017). Does gratitude enhance prosociality? A meta-analytic review. *Psychological Bulletin*. https://doi.org/10.1037/bulo 000 103

Martin, F., & Farnum, J. (2002). Animal-assisted therapy for children with pervasive developmental disorders. *Western Journal of Nursing Research*, *24*(6), 657–670.

Menting, A. T. A., Castro, B. O. de, & Matthys, W. (2013). Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behavior: A meta-analytic review. *Clinical Psychology Review*, *33*(8), 901–913. https://doi.org/10.1016/j.cpr.2013.07.006

Miller, S. C., Kennedy, C., DeVoe, D., Hickey, M., Nelson, T., & Kogan, L. (2009). An examination of changes in oxytocin levels in men and women before and after interaction with a bonded dog. *Examination of Changes in Oxytocin Levels in Men and Women before and after Interaction with a Bonded Dog*, *22*(1), 31–42.

Nagasawa, M., Kikusui, T., Onaka, T., & Ohta, M. (2009). Dog's gaze at its owner increases owner's urinary oxytocin during social interaction. *Hormones and Behavior*, *55*(3), 434–441. https://doi.org/10.1016/j.yhbeh.2008.12.002

Newton, E. K., Laible, D., Carlo, G., Steele, J. S., & McGinley, M. (2014). Do sensitive parents foster kind children, or vice versa? Bidirectional influences between children's prosocial behavior and parental sensitivity. *Developmental Psychology*, *50*(6), 1808–1816. https://doi.org/10.1037/a0036495

Ogden, T., & Hagen, K. A. (2008). Treatment effectiveness of parent management training in Norway: A randomized controlled trial of children with conduct problems. *Journal of Consulting and Clinical Psychology*, *76*(4), 607–621. https://doi.org/10.1037/0022-006X.76.4.607

Piek, J. P., Kane, R., Rigoli, D., Mclaren, S., Roberts, C. M., Rooney, R., … Straker, L. (2015). Does the Animal Fun program improve social-emotional and behavioural outcomes in children aged 4-6 years? *Human Movement Science*, *43*, 155–163.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2015). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1 – 113. Retrieved from http://CRAN.R-project.org/package=nlme

R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Smallwood, K., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly*, *23*, 10–26.

Revelle, W. (2014). *psych: Procedures for Personality and Psychological Research*. Evanston, Il: Northwestern University. Retrieved from http://cran.r-project.org/web/packages/psych/.

Rochat, P., Dias, M. D. G., Liping, G., Broesch, T., Passos-Ferreira, C., Winning, A., & Berg, B. (2009). Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *Journal of Cross-Cultural Psychology*, *40*(3), 416–442. https://doi.org/10.1177/0022022109332844

Roth-Hanania, R., Davidov, M., & Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: Early signs of concern for others. *Infant Behavior and Development*, *34*(3), 447–458. https://doi.org/10.1016/j.infbeh.2011.04.007

Samuels, W. E. (2018). *Support for the valid use of a Chinese version of the Teacher Observation of Classroom Adaptation–Checklist (TOCA-C) Prosociality and Disruptiveness subscales in elementary schools*. Manuscript in preparation.

Samuels, W. E., Meers, L., & Normando, S. (2016). Improving upper elementary students' humane attitudes and prosocial behaviors through an in-class humane education program. *Anthrozoös*, *29*(4), 597–610. https://doi.org/10.1080/08927936.2016.1228751

Samuels, W. E., & Reinhartz, J. (2000). Teen voices: Insight for fostering prosocial behaviors in school. *Teacher Educator's Journal*, *10*(1), 11–27.

Schonert-Reichl, K. (2012). Promoting empathy in school-aged children current state of the field and implications for research and practice. In *School rampage shootings and other youth disturbances: Early preventative interventions* (pp. 159–203).

Schonert-Reichl, K. A., Smith, V., Zaidman-Zait, A., & Hertzman, C. (2012). Promoting children's prosocial behaviors in school: Impact of the "Roots of Empathy" program on the social and emotional competence of school-aged children. *School Mental Health*, *4*(1), 1–21. https://doi.org/10.1037/t01093-000

Schonert-Reichl, K., Oberle, E., Lawlor, M. S., Abbott, D., Thomson, K., Oberlander, T. F., & Diamond, A. (2015). Enhancing cognitive and social-emotional development through a simple-to-administer

mindfulness-based school program for elementary school children: A randomized controlled trial. *Developmental Psychology*, *51*(1), 52–66. https://doi.org/10.1037/a0038454

Serpell, J. (1999). Animals in children's lives. *Society & Animals*, *7*(2), 87.

Sheidow, A. J., Gorman-Smith, D., Tolan, P. H., & Henry, D. B. (2001). Family and community characteristics: Risk factors for violence exposure in inner-city youth. *Journal of Community Psychology*, *29*(3), 345–360.

Shih, M. (1999). Culture, Literacy, and Learning English: Voices From the Chinese Classroom. *TESOL Quarterly*, *33*(1), 159–161.

Singer, J. D., & Willet, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. New York: Oxford University Press.

Sommerville, J. A., Schmidt, M. F. H., Yun, J., & Burns, M. (2013). The development of fairness expectations and prosocial behavior in the second year of life. *Infancy*, *18*(1), 40–66. https://doi.org/10.1111/j.1532-7078.2012.00129.x.

Sorrentino, R. M., Szeto, A. C. H., Chen, X., & Wang, Z. (2013). Uncertainty regulation across cultures: An exploration of individual differences in Chinese and Canadian children in the classroom. *Personality and Individual Differences*, *54*(3), 378–382. https://doi.org/10.1016/j.paid.2012.10.005

Sprinkle, J. E. (2008). Animals, empathy, and violence: Can animals be used to convey principles of prosocial behavior to children? *Youth Violence and Juvenile Justice*, *6*(1), 47–58. https://doi.org/10.1177/1541204007305525

Staub, E. (1971). Helping a person in distress: The influence of implicit and explicit "rules" of conduct on children and adults. *Journal of Personality and Social Psychology*, *17*(2), 137–144.

Stoolmiller, M., Eddy, J. M., & Reid, J. B. (2000). Detecting and describing preventive intervention effects in a universal school-based randomized trial targeting delinquent and violent behavior.

Journal of Consulting and Clinical Psychology, *68*(2), 296–306. https://doi.org/10.1037/0022-006X.68.2.296

Strayer, J., & Roberts, W. (2004). Children's anger, emotional expressiveness, and empathy: Relations with parents' empathy, emotional expressiveness, and parenting practices. *Social Development*, *13*(2), 229–254. https://doi.org/10.1111/j.1467-9507.2004.000265.x

Stürmer, S. (2009). The psychological study of group processes and intergroup relations in prosocial behavior. In S. Stürmer & M. Snyder (Eds.) *The Psychology of Prosocial Behavior* (pp. 1–10). Oxford, UK: Wiley-Blackwell. Retrieved from http://site.ebrary.com/lib/csi/detail.action?docID=10343025

Tasimi, A., & Young, L. (2016). Memories of good deeds past: The reinforcing power of prosocial behavior in children. *Journal of Experimental Child Psychology*, *147*, 159–166. https://doi.org/10.1016/j.jecp.2016.03.001

Taylor, Z. E., Eisenberg, N., Spinrad, T. L., Eggum, N. D., & Sulik, M. J. (2013). The relations of ego-resiliency and emotion socialization to the development of empathy and prosocial behavior across early childhood. *Emotion*, *13*(5), 822–831. https://doi.org/10.1037/a0032894

Thielmann, I., & Böhm, R. (2016). Who does (not) participate in intergroup conflict? *Social Psychological and Personality Science*, *7*(8), 778–787. https://doi.org/10.1177/1948550616660160

Thompson, R. A., & Newton, E. K. (2013). Baby altruists? Examining the complexity of prosocial motivation in young children. *Infancy*, *18*(1), 120–133. https://doi.org/10.1111/j.1532-7078.2012.00139.x.

Unti, B. & DeRosa, B. (2003). Humane education: Past, present, and future. In D. J. Salem & A. N. Rowam (Eds.), *The State of the Animals II: 2003*. Washington: Humane Society Press, 27-50.

Unti, B. & Rowan, A. (2001). A social history of animal protection in the post-World War II period. In D. Salem & A. Rowan (Eds.), *The State of Animals: 2001*. Washington: Humane Society Press, 21 – 37.

Usami, S. (2014). A convenient method and numerical tables for sample size determination in longitudinal-experimental research using multilevel models. *Behavior Research Methods*, *46*(4), 1207–1219. https://doi.org/10.3758/s13428-013-0432-0

Wandersman, A., & Nation, M. (1998). Urban neighborhoods and mental health: psychological contributions to understanding toxicity, resilience, and interventions.(Urban Life: Adversity and Reward). *The American Psychologist*, *53*(6), 647–656. https://doi.org/10.1037/0003-066X.53.6.647

Wentzel, K. R., Barry, C. M., & Caldwell, K. A. (2004). Friendships in middle school: Influences on motivation and school adjustment. *Journal of Educational Psychology*, *96*(2), 195–203. https://doi.org/10.1037/0022-0663.96.2.195

Williams, A., O'Driscoll, K., & Moore, C. (2014). The influence of empathic concern on prosocial behaviour in children. *Frontiers in Psychology*, *5*, 8 pp.

Windhager, S., Atzwanger, K., Bookstein, F. L., & Schaefer, K. (2011). Fish in a mall aquarium: An ethological investigation of biophilia. *Landscape and Urban Planning*, *99*(1), 23–30.

Yoo, H., Feng, X., & Day, R. (2013). Adolescents' empathy and prosocial behavior in the family context: A longitudinal study. *A Multidisciplinary Research Publication*, *42*(12), 1858–1872. https://doi.org/10.1007/s10964-012-9900-6

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, *28*(1), 126–136.

Zhu, Y., Guan, X., & Li, Y. (2015). The effects of intergroup competition on prosocial behaviors in young children: A comparison of 2.5–3.5 year-olds with 5.5–6.5 year-olds. *Frontiers in Behavioral Neuroscience*, *9*, 16 pp.