Abstract Title Page

Title:
Replication of effects of the *Positive Action* program in randomized trials in Hawai'i and Chicago schools

Authors and Affiliations:
Brian R. Flay, D.Phil.,
Professor, Health Promotion Health Behavior, School of Social and Behavioral Health Sciences, College of Public Health and Human Sciences, Oregon State University

Acknowledgements:
The Hawai'i trial was funded by the National Institutes on Drug Abuse of the National Institutes of Health (DA13474 and DA018760).

The Chicago trial (ClinicalTrials.gov #NCT01025674) was funded by grants from the Institute of Education Sciences (IES), US Department of Education: R305L030072, R305L030004 and R305A080253 to the University of Illinois, Chicago (2003-05) and Oregon State University (2005-12). The initial phase (R305L030072), a component of the Social and Character Development (SACD) Research Consortium, was a collaboration among IES**,** the Centers for Disease Control and Prevention's (CDC) Division of Violence Prevention, Mathematica Policy Research Inc. (MPR), and awardees of SACD cooperative agreements (Children's Institute, New York University, Oregon State University, University at Buffalo-SUNY, University of Maryland, University of North Carolina-Chapel Hill, and Vanderbilt University).

The SACD research program includes multi-program evaluation data collected by MPR and complementary research study data collected by each grantee. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Institute of Education Sciences, CDC, MPR, or every Consortium member, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. Correspondence concerning this article should be addressed to the Author.

We are extremely grateful to the participating Chicago Public Schools (CPS), their principals, teachers, students, and parents. We thank the CPS Research Review Board and Office of Specialized Services, especially Drs. Renee Grant-Mitchell and Inez Drummond, for their invaluable support of this research. I am also extremely grateful to my co-investigators on these trials and the co-authors on the many papers from them.

Notice of potential conflict of interest:
The research described herein was conducted using the program, the training, and technical support of *Positive Action*, Inc. in which Brian Flay's spouse holds a significant financial interest. Issues regarding conflict of interest were reported to the relevant institutions and appropriately managed following the institutional guidelines.

Background / Context:

Several social-emotional learning (SEL) or social-emotional and character development (SECD) programs have been shown to be effective at improving SEL/SECD skills, and some have also provided evidence of effectiveness in improving student behavior and academic achievement (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). Very few SEL or SECD programs have replicated effects for all three of these domains of outcomes with students of different cultural/ethnic backgrounds in schools of different socioeconomic status and urbinicity. The Society for Prevention Research, among others, has emphasized the importance of replication of findings (Flay et al., 2005; Valentine et al., 2011). In education research, replication with different populations in different settings is important for understanding the generalizability of findings. In this paper, we report replication of effects of an SEL/SECD program across two very different groups of students and contexts.

Purpose / Objective / Research Question / Focus of Study:

This paper reports results from two cluster-randomized trials of the *Positive Action (PA)* program in elementary schools in two disparate locations. Program effects are presented on theoretically predicted outcomes of school quality and student character, negative behaviors, emotional/mental health, positive health behaviors, and school performance.

Setting:

One study was conducted in 2002-2006 in 20 suburban and rural schools on three Hawai'i islands. The second was conducted in 2004-2010 in 14 high-poverty, inner-city Chicago schools.

Population / Participants / Subjects:

In Hawai'i students were approximately 26% Hawaiian or part-Hawaiian, 9% White, 21% Asian, 23% multi-ethnic, 5% other Pacific Islander, and 16% Other or unknown(Beets et al., 2009), about 55% received free or reduced price lunches (Snyder et al., 2010) (total N = 1784), and were followed from grades 2/3 to grades 5/6. In Chicago, students were approximately 54% African American, 31% Hispanic, 8% Caucasian, 4% Asian-American and 3% Other or unknown, 90% received free or reduced-price lunches (Lewis, Schure, et al., 2013) and were followed from grade 3 to grade 8.

Intervention / Program / Practice:

As evaluated, *Positive Action* is a school-based program that includes school-wide climate change and a detailed curriculum with lessons 2-4 times a week—approximately 140 15-minute lessons per grade K-8 and 82 15-20 minute lessons per grade 7 and 8. Lessons for each grade level are scripted and age-appropriate and have all the materials to teach the lesson including posters, puppets, music, games, and other hands–on materials integrated into the lessons. The students' materials come prepared for a classroom of 30 students and include activity booklets, journals and other lesson aids. The content of the program is included in six units that form the foundation for the whole program. The first unit teaches the philosophy of the program and the Thoughts-Actions-Feelings about Self Circle, and provides an introduction to the nature and relevancy of positive and negative actions/behaviors. Units 2-6 teach the positive actions for the physical, intellectual, social and emotional areas. There are two school-wide climate development kits (elementary and secondary) and a Counselor's Kit. The contents delivered through the climate development and counselor kits reinforce the classroom curriculum through coordinating the efforts of the entire school in the practice and reinforcement of positive actions. [Family and community components of the program were not used in these trials.]

Research Design:

In both trials, schools whose principals agreed to participate in the study were formed

into matched pairs using school-level demographic, behavioral and achievement data, and then randomly assigned to program or wait-listed control (business as usual) (Ji, DuBois, Flay, & Brechling, 2008). In Hawai'i, 10 pairs of schools were included; but in Chicago, only 7 pairs were included (because of requirements of the IES-sponsored multi-site Social and Character Development (SACD) study of which the Chicago trial was a part).

Program implementation data were collected from teachers at the end of each of the 6 units of *PA* and the end of each year, and from students at the end of each year (Beets et al., 2008; Malloy et al., under review). At the end of each year, teachers and students in both program and control (C) schools also responded to questions about use of SACD-type strategies (Beets & Flay, 2007; Social and Character Development Research Consortium, 2010).

Outcome data were primarily student self-reports of behavior and school-level archival data on disciplinary referrals/suspensions and achievement (standardized test scores). In the Hawai'i trial, assessment was limited to one class period each time, during which survey data were collected on a social-emotional and character development scale (Ji, DuBois, & Flay, in press). In the Chicago trial, assessment took place over two class periods, during which data were collected on multiple behavioral outcomes (e.g., empathy, altruism, depression, anxiety, disruptive behavior, bullying) as well as the same social-emotional and character development scale used in Hawai'i. At grade 5, items were added to assess substance use and violence in both the Hawai'i and Chicago trials, and also sexual behaviors in the Hawai'i trial.

For both trials, school-level archival data were obtained from a year or more before the trial to a year or more after the trial regarding disciplinary referrals or suspensions and standardized test scores. For the Hawai'i trial, we were also able to obtain school-level data on student, parent and teacher responses to a school quality survey that the school district administers every two years.

Data Collection and Analysis:

In Hawai'i, students in half of the schools were surveyed in Spring of the year prior to the start of the trial (at the end of grades 1 or 2) and students in the other half were surveyed in Fall of the first year of the trial (at the beginning of grades 2 or 3). Follow-up surveys were administered to these cohorts of students at the end of the 4 subsequent school years, for a total of 5 waves of data. In Chicago, the selected cohort of students was surveyed at the beginning and end of grade 3, the beginning and end of grade 4 and the end of grade 5. A second round of funding then allowed for follow-up surveys to be administered at the beginning and end of grade 7 and the end of grade 8, for a total of 8 waves of data.

Given the high-risk nature of the selected schools in both trials, student mobility was high; students who left study schools during the study were not followed and students who entered project schools were added to the study. Statistical analysis using multi-level latent growth curve modeling took account of the missing data (after leavers left and before joiners joined) (Vuchinich, Flay, Aber, & Bickman, 2012). We report modeled standardized effect sizes (Hedge's *g*) (Hedges & Olkin, 1985) at endpoints; all reported effects are statistically significant at $p < .05$, 2-tailed, unless characterized as marginal. Potential moderation by student mobility pattern and gender were tested (ethnicity was not tested because of confounding with school).

Findings / Results:

In both trials, there were no significant baseline differences between conditions on the school-level matching variables (Beets et al., 2009; Li et al., 2011). In Hawai'i, there were also no differences on student pretest data. In Chicago, of over 50 student-level measures, baseline differences were significant for only four, and two of these favored the C group and two favored

the program group (Lewis, DuBois, et al., under review).

Program implementation was probably higher on average in Hawai'i (Beets et al., 2008) than in Chicago (Malloy et al., under review). Table 1 summarizes program effects.

School quality: School-level mean scores on Hawai'i student, parent and teacher reports of school quality improved significantly more in *PA* schools than C schools (ESs of 1.31, 1.26 and 1.61) (Snyder, Vuchinich, Acock, Washburn, & Flay, 2012). In Chicago, students in PA schools reported stronger teacher (ES=1.26) and school attachment (ES=1.25) than students in C schools, less victimization (ES=-.11), more positive school orientation (ES=.38) and school climate (ES=.39) and better perceptions of the neighborhood (ES=.23) (paper in preparation).

Student character: Scores on the SECD scale items decreased as students got older in both Hawai'i and Chicago, but the decline was mitigated in the *PA* condition of both trials (ESs .46 and .58, respectively) (Washburn et al., 2011). Data from a third small CRCT (4 schools per condition) in a rural southeastern location showed a similar, though weaker, pattern with an ES of .22 (Washburn et al., 2011). Effects were moderated by gender occurred for three subscales of SECD; namely, program effects on honesty, self-control and respect for teachers were stronger for girls than boys (Lewis, Vuchinich, et al., under review).

Negative behaviors: The negative behaviors that were first assessed at grade 5 were all reduced in *PA* schools compared with C schools in both trials (Beets et al., 2008; Lewis et al., 2012; Lewis, Schure, et al., 2013; Li et al., 2011). For substance use, the effect sizes were -1.05 at grade 5 in Hawai'i and -.35 and -.27 at grades 5 and 8, respectively in Chicago. For extreme violence, the effect sizes were -1.39 at grade 5 in Hawai'i and -.26 and -.54 at grades 5 and 8, in Chicago. Hawai'i students also demonstrated reduced early onset of sexual behavior (ES=-1.10). Chicago students also demonstrated reduced bullying (ES=-.26 and -.39 at grades 5 and 8, respectively) and disruptive behaviors (ES=-.23 and -.50 at grades 5 and 8, respectively). Archival data from both Hawai'i (Snyder et al., 2010) and Chicago (Lewis, Schure, et al., 2013) demonstrated program effects on disciplinary suspensions, with ESs of -.87 and -.27.

Other social-emotional outcomes: Significant effects were also demonstrated in the Chicago trial for a range of other outcomes, for example, improvement in affiliation with deviant peers (ES=-.45) or positive peers (ES=.46), aggressive problem solving (ES=-.76), empathy ES=.26) and altruism (ES=.21) (Lewis, Vuchinich, et al., under review); reductions in depression (ES=-.14) anxiety (ES=-.26) and improvements in life satisfaction (Lewis, DuBois, et al., 2013).

Positive health-related behaviors: Significant improvements were also demonstrated in the Chicago trial on student self-reported hygiene (ES=.48) and consumption of unhealthy food (ES=-.19), with marginal effects on healthy diet and exercise (ES=.21), and unhealthy BMI z-scores (ES=-.21) (Bavarian et al., under review).

School performance: School-level student absenteeism rates were reduced in both Hawai'i (Snyder et al., 2010) and Chicago (Bavarian et al., 2013) (ESs of -.65 and -.78).

From Hawai'i, school-level standardized test scores were available for state-level (Hawai'i Content and Performance Standards - HCPS) as well nationally-standardized (Terra Nova, 2nd Edition) tests. Growth-curve models demonstrated significant effects for both reading and math on both the HCPS (ESs of .65 and 1.1, respectively) and the Terra Nova (ES of .54 and .52, respectively) (Snyder et al., 2010).

In Chicago, significant program effects were also observed for student self-reports of disaffection with learning (ES=-.19) and teacher ratings of student academic motivation (ES=.39) and performance (ES=.19) (Bavarian et al., 2013). Effects on school-level academic outcomes from the Chicago trial were not as strong as from the Hawai'i trial, but some

interesting patterns emerged (Bavarian et al., 2013). There were marginally significant effects on math performance of all grades 3-8 students combined (ES=.38). Gender or ethnicity did not moderate these effects. Although the overall effect (for grades 3-8 combined) on reading was not significant (ES=.22), there were marginal effects for boys (ES=.33), large and significant effects for cohort students transitioning between grades 7 and 8 (that is, the value-added score, ES=.83) and for African American students (ES=.50), especially African American boys (ES=1.50).

Conclusions:

Consistent with the presumed theory of action of the *Positive Action* program, both trials found strong effects on indicators of school quality/climate. The fact that the measures were school-level means from student, parent and teacher surveys in Hawai'i (collected independently from the research trial) and student self-reports in Chicago make the consistency of these effects all the more remarkable.

Students from both sites also reported a slower decline in social-emotional and character development (SECD) skills. Data from a third small RCT in a rural southeastern location, conducted independently of these researchers and the *PA* developer showed a similar pattern.

Negative behaviors were significantly reduced in *PA* schools compare with C schools in both trials using the same measures of substance use and serious violence. A similar pattern of effects was observed for additional measures of disruptive and bullying behaviors used in the Chicago trial. In addition, the Chicago trial produced evidence of positive program effects on a range of other externalizing and internalizing behaviors. Very few of the effects were moderated by gender (for which effects were stronger for girls), and none by mobility patterns (suggesting equally strong program effects for stayers, leavers and joiners).

Very strong effects of the *PA* program on academic outcomes were observed in the Hawai'i trial. Although the effect sizes were not as large in the Chicago trial, there were some very important effects, including a large effect on a value-added metric for students transitioning between grades 7 to 8, and very large effects for African American students, especially boys. The latter suggest an important role for SEL/SECD in general, and the *PA* program in particular, for helping to close the minority gap in test scores.

Overall, major findings from Hawai'i were replicated in Chicago, although often with smaller effect sizes. We attribute this to the more challenging circumstances of disadvantaged students in inner-city schools. The Hawai'i trial was conducted with students of many ethnic backgrounds in a range of suburban and rural schools on three islands. The Chicago trial was conducted with primarily African American, and Hispanic students in poor, inner-city neighborhoods. Also, program implementation was probably higher in Hawai'i than in Chicago.

Some limitations to these trials need to be acknowledged. Both were small for cluster-randomized trials. Nevertheless, careful matching and analysis of multiple waves of data appears to have provided sufficient statistical power to detect many effects. Most of the results rely on student self-reports, and these could be biased in socially desirable directions. However, the existence of school-level data collected independently of these trials, especially about school climate in Hawai'i and student performance in both trials, mitigates this concern.

The fact that the basic pattern of findings was replicated across such diverse contexts suggests that these findings are likely generalizable across a wide range of settings. The *PA* developer claims that the program is universal, and uses the fact that it has been implemented in over 15,000 U.S. schools and in many thousands of schools in many other countries to support this claim. The results from these randomized trials in two very different settings, together with results from prior quasi-experimental evaluations, provide further support to this claim.

Appendices

Appendix A: References

Bavarian, N., Lewis, K. M., Acock, A., DuBois, D. L., Zi, Y., Vuchinich, S., . . . Flay, B. R. (under review). *Direct and Mediated Effects of a Social-Emotional Learning and Health Promotion Program on Adolescent Health Outcomes: A Matched-Pair, Cluster-Randomized Controlled Trial*. Oregon State University.

Bavarian, N., Lewis, K. M., DuBois, D. L., Acock, A., Vuchinich, S., N., S., . . . Flay, B. R. (2013). Using Social-Emotional and Character Development to Improve Academic Outcomes: A Matched-Pair, Cluster-Randomized Controlled Trial in Low-Income, Urban Schools. *Journal of School Health, 83*(11), 771-779.

Beets, M. W., & Flay, B. R. (2007). Teacher- and school-specific factors related to implementation of the Positive Action program. *Research Quarterly for Exercise and Sport, 78*(1), A24-A24.

Beets, M. W., Flay, B. R., Vuchinich, S., Acock, A., Li, K. K., & Allred, C. G. (2008). School Climate and Teachers' Beliefs and Attitudes Associated with Implementation of the Positive Action Program: A Diffusion of Innovations Model. *Prevention Science, 9*(4), 264-275. doi: DOI 10.1007/s11121-008-0100-2

Beets, M. W., Flay, B. R., Vuchinich, S., Snyder, F. J., Acock, A., Li, K.-K., . . . Durlak, J. A. (2009). Use of a social and character development program to prevent substance use, violent behaviors, and sexual activity among elementary students in Hawai'i. *American Journal of Public Health, 99*(8), 1438-1445. doi: 10.2105/AJPH.2008.142919

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*(1), 405-432.

Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., . . . Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6*(3), 151-175. doi: DOI 10.1007/s11121-005-5553-y

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* San Diego, CA: Academic Press.

Ji, P., DuBois, D. L., & Flay, B. R. (in press). Social-Emotional and Character Development Scale: Development and Initial Validation with Urban Elementary School Students. *Journal of Research on Character Education*.

Ji, P., DuBois, D. L., Flay, B. R., & Brechling, V. (2008). "Congratulations, you have been randomized into the control group!(?)": Issues to consider when recruiting schools for matched-pair randomized control trials of prevention programs. *Journal of School Health, 78*(3), 131-139.

Lewis, K. M., Bavarian, N., Acock, A., DuBois, D. L., Ji, P., Schure, M. B., . . . Flay, B. R. (2012). Impact and Mediators of a Social-Emotional and Character Development Program on Adolescent Substance Use. *International Journal of Emotional Education, 4*(1), 56-78.

Lewis, K. M., DuBois, D. L., Bavarian, N., Acock, A., Silverthorn, N., Day, J., . . . Flay, B. R. (2013). Effects of Positive Action on the Emotional Health of Urban Youth: A Cluster-Randomized Trial. *Journal of Adolescent Health*. doi: doi: 10.1016/j.jadohealth.2013.06.012

Lewis, K. M., DuBois, D. L., Ji, P., Day, J., Silverthorn, N., Bavarian, N., . . . Flay, B. R. (under review). Meeting the challenges of longitudinal cluster-based trials in schools: Lessons from the Chicago trial of *Positive Action*

Lewis, K. M., Schure, M. B., Bavarian, N., DuBois, D. L., Day, J., Ji, P., . . . Flay, B. R. (2013). Problem Behavior and Urban, Low-Income Youth: A Randomized Controlled Trial of Positive Action in Chicago. *American Journal of Preventive Medicine, 44*(6), 622-630. doi: 10.1016/j.amepre.2013.01.030

Lewis, K. M., Vuchinich, S., Ji, P., DuBois, D. L., Acock, A., Bavarian, N., . . . Flay, B. R. (under review). Effects of the *Positive Action* Program on Indicators of Positive Youth Development Among Urban Youth.

Li, K.-K., Washburn, I., DuBois, D. L., Vuchinich, S., Ji, P., Brechling, V., . . . Flay, B. R. (2011). Effects of the *Positive Action* programme on problem behaviors in elementary school students: A matched-pair, randomized control trial in Chicago. *Psychology & Health, 26*(2), 187-204. doi: 10.1080/08870446.2011.531574

Malloy, M., Acock, A., DuBois, D. L., Vuchinich, S., SIlverthorn, N., Ji, P., & Flay, B. R. (under review). Teachers' Perceptions of School Organizational Climate Characteristics as Predictors of the Implementation Dosage and Quality of a Social and Emotional Character Development Program.

Snyder, F. J., Flay, B. R., Vuchinich, S., Acock, A., Washburn, I. J., Beets, M. W., & Li, K.-K. (2010). Impact of a social-emotional and character education program on school-level indicators of academic achievement, absenteeism, and disciplinary outcomes: A matched-pair, cluster randomized, controlled trial. *Journal of Research on Educational Effectiveness, 3*(1), 26-55.

Snyder, F. J., Vuchinich, S., Acock, A., Washburn, I. J., & Flay, B. R. (2012). Improving elementary school quality through the use of a social-emotional and character development program: A matched-pair, cluster-randomized, controlled trial in Hawai'i. *Journal of School Health, 82*(1), 11-20. doi: Doi 10.1111/J.1746-1561.2011.00662.X

Social and Character Development Research Consortium. (2010). Efficacy of school-wide programs to promote social and character development and reduce problem behavior in elementary school children. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Valentine, J. C., Biglan, A., Boruch, R. F., Collins, L. M., Flay, B. R., Kellam, S., . . . Schinke, S. P. (2011). Replication in Prevention Science. *Prevention Science, 12*, 103-117. doi: 10.1007/s11121-011-0217-6

Vuchinich, S., Flay, B. R., Aber, L., & Bickman, L. (2012). Person mobility in the design and analysis of cluster-randomized cohort prevention trials. *Prevention Science, 13*(3), 300-313. doi: 10.1007/s11121-011-0265-y

Washburn, I. J., Acock, A. C., Vuchinich, S., Snyder, F. J., Li, K.-K., Ji, P., . . . Flay, B. R. (2011). Effects of a social-emotional and character development program on the trajectory of behaviors associated with character development: Findings from three randomized trials. . *Prevention Science, 12*, 314-323. doi: 10.1007/s11121-011-0230-9

Appendix B:

Table 1: Summary of Effects of *Positive Action* from Randomized Trials

Hawai'i and Southeastern results are end of grade 5; Chicago results are end of grade 8

%II = Improvement Index; *%RI = % relative improvement (of PA group relative to control group)*

Outcome Domain/Measure	Site (grade)	ES	%II or *%RI*	p-value*
Social climates/environments				
Teacher Attachment Scale	Chicago (8th)	1.26	39.62%	0.001
School Attachment Scale	Chicago (8th)	1.25	39.44%	0.001
Teacher Rewards for Prosocial Behavior	Chicago (8th)	0.79	28.52%	0.001
Parent Rewards for Prosocial Behavior	Chicago (8th)	0.84	29.95%	0.001
Positive School Orientation	Chicago (8th)	0.38	14.80%	0.001
Inventory of School Climate Scale	Chicago (8th)	0.39	15.17%	0.05
Student Perception of Neighborhood Context	Chicago (8th)	0.23	9.10%	0.05
Victimization Scale	Chicago (8th)	-0.11	*-22.12%*	0.05
School Quality				
Teacher reports	Hawai'i (5th)	1.61	*21.00%*	0.05
Parent reports	Hawai'i (5th)	1.26	*13.00%*	0.05
Student reports	Hawai'i (5th)	1.31	*16.00%*	0.05
Social-emotional and character				
SECD – General Character (Composite)	SE state (5th)	0.41	15.90%	0.001
	Hawai'i (5th)	0.46	17.72%	0.001
	Chicago (8th)	0.58	21.90%	0.001
SECD - Prosocial Interaction	Chicago (8th)	0.49	18.79%	0.001
SECD –Self Control	Chicago (8th)	0.60	22.57%	0.001
SECD – Honesty**	Chicago (8th)	0.45	17.36%	0.05
SECD – Self-Development**	Chicago (8th)	0.32	12.55%	0.05
SECD – Respect for teacher**	Chicago (8th)	0.76	27.64%	0.001
SECD – Respect for parent	Chicago (8th)	0.65	24.22%	0.001
Self-concept (PAI/Positive Feelings Scale)	Chicago (8th)	0.32	12.55%	0.10
Positive Peer Affiliation	Chicago (8th)	0.46	17.72%	0.001
Deviant Peer Affiliation	Chicago (8th)	-0.45	*-25.86%*	0.001
Aggressive Problem Solving	Chicago (8th)	-0.76	*-80.95%*	0.001
Competent Problem Solving	Chicago (8th)	0.05	*0.21%*	ns
Positive Morality	Chicago (8th)	0.13	*0.25%*	ns
Negative Morality	Chicago (8th)	-0.57	*-26.96%*	0.01
Children's Empathy Scale	Chicago (8th)	0.26	10.26%	0.10
Altruism Scale	Chicago (8th)	0.21	8.32%	0.05
Emotional/Mental Health				
BASC Depression Scale	Chicago (8th)	-0.14	*-17.16%*	0.05
BASC Anxiety Scale	Chicago (8th)	-0.26	*-18.34%*	0.00
Positive Affect (Scale for Children)	Chicago (8th)	0.17	6.75%	0.10
Life Satisfaction	Chicago (8th)	0.13	5.17%	0.05
Self-esteem				
Peer self-esteem	Chicago (8th)	0.37	14.43%	0.05
School self- esteem	Chicago (8th)	0.46	17.72%	0.01

SE Formation and Maintenance - Adaptive	Chicago (8th)	0.31	12.17%	0.05
SE Motivation Scale	Chicago (8th)	0.47	18.08%	0.05
Health Behaviors				
Hygiene Scale	Chicago (8th)	0.48	18.40%	0.05
Unhealthy Food	Chicago (8th)	-0.19	-7.50%	0.05
Healthy Food and Exercise	Chicago (8th)	0.21	8.32%	0.10
Sleep	Chicago (8th)	0.35	13.70%	0.21
Unhealthy Body Mass Index Percentile	Chicago (8th)	-0.21	*-23.16%*	0.10
Negative Behaviors				
Violence				
Normative Beliefs Supporting Aggression	Chicago (8th)	-0.68	*-59.26%*	0.01
Aggression Scale (Bullying)	Chicago (5th)	-0.59	*-41.00%*	0.05
	Chicago (8th)	-0.39	*-51.13%*	0.01
Parent Report of Bullying	Chicago (8th)	-0.31	*-12.81%*	0.05
Serious violence	Hawai'i (5th)	-1.39	*-51.85%*	0.01
	Chicago (5th)	-0.35	*-36.00%*	0.05
	Chicago (8th)	-0.54	*-41.79%*	0.05
Violence - teacher report	Hawai'i (5th)	-0.91	*-34.32%*	0.05
Disruptive Behaviors	Chicago (5th)	-0.47	*-27.00%*	ns
	Chicago (8th)	-0.50	*-71.79%*	0.001
Substance use				
Substance Use (Composite)	Hawai'i (5th)	-1.05	*-44.23%*	0.01
	Chicago (5th)	-0.41	*-31.00%*	0.05
	Chicago (8th)	-0.27	*-10.67%*	0.01
Substance use - teacher report	Hawai'i (5th)	-0.72	*-47.67%*	0.10
Cigarette Smoking	Hawai'i (5th)	-0.44	*-47.37%*	0.05
	Chicago (8th)	-0.21	*-31.11%*	0.05
Alcohol Use	Hawai'i (5th)	-0.44	*-46.28%*	0.05
	Chicago (8th)	-0.35	*-28.04%*	0.05
Gotten Drunk	Hawai'i (5th)	-0.75	*-69.81%*	0.05
	Chicago (8th)	-0.29	*-40.51%*	0.01
Illegal drug use	Hawai'i (5th)	-0.82	*-73.17%*	0.05
Marijuana Use	Chicago (8th)	-0.23	*-37.03%*	0.05
Gotten high on drugs	Hawai'i (5th)	-0.99	*-80.00%*	0.05
Sexual activity	Hawai'i (5th)	-1.10	*-82.61%*	0.10
Disciplinary Actions (School level)				
Referrals	Chicago (8th)	-0.58	*-46.19%*	0.01
Suspensions	Hawai'i (5th)	-0.96	*-69.40%*	0.10
	Chicago (8th)	-0.27	*-23.56%*	0.01
Academic Outcomes				
Student-level data				
Teacher rating academic motivation	Chicago (8th)	0.39	15.17%	0.05
Disaffection with Learning (4 items)	Chicago (8th)	-0.19	-7.53%	0.01
Teacher rating academic ability	Chicago (8th)	0.14	5.57%	0.05
School-level data				
Absenteeism	Hawai'i (5th)	-0.63	-23.60%	0.00

	Chicago (8th)	-0.78	-28.23%	0.05
Retained in grade	Hawai'i (5th)	-0.84	*-60.00%*	ns
Reading Standardized Scores				
All students	Hawai'i (5th)	0.58	21.90%	0.10
Reading State Test Scores				
All students	Hawai'i (5th)	0.72	26.42%	0.05
Value-added metric at 8th grade	Chicago (8th)	0.83	29.67%	0.05
All students grades 3-8	Chicago	0.22	8.71%	0.32
Males	Chicago	0.33	12.93%	0.20
African Americans	Chicago	0.50	19.15%	0.20
African American Males	Chicago	1.50	43.32%	0.05
Math Standardized Test Scores				
All students	Hawai'i (5th)	0.50	19.15%	ns
Math State Test Scores				
All Students	Hawai'i (5th)	0.69	25.49%	0.05
All students grades 3-8	Chicago	0.38	14.80%	0.15

Notes:

* Some analyses used confidence intervals, so all p-values simply reported here as < a critical value.

** See text for gender moderation of these outcomes.