Study 4

Grade Level Effects of the Incredible Years Teacher Training Program on Emotion Regulation and Attention

Authors and Affiliations:

Desiree W. Murray, PhD
Center for Child and Family Policy & Social Science Research Institute, Duke University

David L. Rabiner, PhD
Trinity College of Arts and Sciences & Department of Psychology and Neuroscience, Duke University

Madeline M. Carrig, PhD
Center for Child and Family Policy & Department of Psychology and Neuroscience, Duke University

Abstract Body

Background / Context:
Teachers' use of effective classroom management skills such as praise, proactive teaching strategies, and non-harsh discipline predicts students' social-emotional competence, fosters the development of emotion regulation, and reduces disruptive behaviors (Pianta, LaParo, Payne, Cox, & Bradley, 2002; Webster-Stratton, Reid & Stoolmiller, 2007). Although classroom management has also been shown to increase student engagement in learning (Creemers, 1994; Stringfield, 1994) there has been little controlled research demonstrating the impact of classroom management interventions on student achievement (Department of Education, 2008). Teacher training interventions therefore warrant more rigorous evaluation with assessment of specific academic effects in addition to investigation of social-emotional and behavioral outcomes.

Professional development for teachers has historically been fairly didactic in nature (Garet et al., 2001; Rose & Church, 1998), which is believed to limit effectiveness and translation to teacher practice change (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005). One well-established program with a more active, collaborative training approach is the Incredible Years Teacher Classroom Management Program (IYT). Although positive effects on young children's social-emotional competence have been demonstrated for this program when implemented in combination with other Incredible Years programs (Webster-Stratton, Reid, & Hammond, 2001; 2004; Webster-Stratton, Reid, & Stoolmiller, 2008), the program has not been evaluated as an independent intervention or with regard to academic outcomes for early elementary students.

As with many teacher-directed universal interventions for behavior problems or social-emotional development, there may be differential effects for students with different risk levels (Bierman et al, 2010). In addition, it is possible that social-emotional learning (SEL) interventions may have varying impact based upon students' developmental levels. Typically, it is believed that such interventions may be more effective for younger students, although there may be other contextual factors that affect outcomes. Thus, moderator analyses can be informative in understanding intervention effects.

Objective / Research Question:
In this IES-funded Goal 3 study, we examined the efficacy of IYT for improving (a) classroom climate and teacher management skills in grades K-2; and (b) students' attention, social-emotional competence, and achievement as moderated by grade level and socio-economic status. We hypothesized that students in younger grade levels and those receiving free/reduced lunch would experience the greatest benefit.

Setting:
Eleven schools across three rural school districts in the southeastern U.S. participated. Rural schools were selected due to their lower than average level of school resources and difficulty recruiting and retaining highly qualified teachers, which are factors placing students at greater risk of educational failure (Bacolod, 2007). However, participating schools varied greatly with regard to student race/ethnicity, poverty and achievement levels (ranging from 16-100% free/reduced lunch, 19-80% white, and 41-90% of 3rd graders at grade level in reading).

Participants:
Participants included 97 K-2 teachers and 1276 students. Teachers were primarily female and Caucasian (94.8% and 81.4%, respectively), with an average of 11.25 years of experience. Thirty percent held a master's degree. Class sizes ranged from 16 to 24 students (M = 19, SD = 1.87). Students were 51% male and 47% received free/reduced lunch. The student sample was ethnically diverse (54% Caucasian, 22% Hispanic, and 17% African-American).

Intervention:
IYT is a teacher training program that is part of a comprehensive series of interventions including parent, child, and teacher training components that were designed to prevent and treat aggressive behavior and conduct problems in young children aged 3-8 years. Its approach includes validated training methods such as video-modeling, behavioral rehearsal of key skills through numerous role plays, classroom practice assignments, and teacher goal setting and self-monitoring. IYT is provided in 4-6 monthly full-day workshops (5 in the present study) that cover building positive relationships with students and parents, proactive classroom management strategies, effective use of incentives, "coaching" students social and emotional development, teaching calm-down and problem-solving, and positive discipline techniques such as redirection, ignoring, and time out. Workshops are led by two trained co-leaders with approximately 12-15 teachers in each group.

In the present study, the average number of workshop training hours across the 5 training days for intervention teachers was 34, with an overall attendance rate of 97%. Each teacher also received two brief consultation visits in their classroom (average of 44 minutes total) and regular emails to support implementation of strategies taught. Intervention fidelity was strong, as indicated by >85% of workshop objectives being met and teacher satisfaction ratings >6 on a 1-7 scale, and was supported by regular consultation calls with the intervention developer, Carolyn Webster-Stratton. In addition, the PI is a nationally certified trainer in the program.

Research Design:
This was a randomized, controlled efficacy study using a 4-level (students, teachers, grade levels, schools) cluster-randomized block design with treatment assignment occurring at level 3 (grade level within school). At least one participating grade level per school was assigned to the intervention condition. Across three cohorts, forty-seven teachers were assigned to the intervention condition (of whom 45 participated) with 598 intervention and 560 control students with data available for analysis. Due to varying numbers of teachers within grade levels across schools, distribution of teachers across grade levels was unequal (K=36, 1^{st}=26, 2^{nd}=35).

Data Collection and Analysis:
Following teacher consent and randomization at the beginning of the school year, trained research assistants blind to randomization status observed each teacher's classroom for approximately 2 hours of instructional time and collected teacher ratings on all students for whom parental consent was obtained. These assessments were repeated in the spring, following completion of approximately five months of intervention.

Multilevel modeling in SAS (version 9.2, PROC MIXED) was used to account for nesting of students within teacher, grade level, and school, and to model change over time. Covariates

included free/reduced lunch, sex and race/ethnicity at the student level; teacher years of experience, percent of students in the class on free/reduced lunch and percent of students with below-average social competence at the classroom level; and percent free/reduced lunch at the school level. Hypothesized moderators of the intervention effect included grade level, student free/reduced lunch status, and their interaction.

Measures

Observational change in teacher practices was assessed with the *CLASS* (Pianta & Hamre, 2005) and the *Teacher Coder Impressions Inventory* (TCI; Webster-Stratton, Reid, & Hammond, 2001). The CLASS is a multi-dimensional standardized instrument that has been widely used in early education classrooms and has been associated with variation in students' achievement and social adjustment. Subscales examined for this study include Positive Climate, Negative Climate, and Behavior Management. The TCI was developed to align more closely with the IYT intervention and has been used primarily in research on the efficacy of IYT. Analyses were conducted on the Competent subscale of this measure. Reliability of both measures was strong (overall kappa for CLASS=.87; TCI=.90).

Student social-emotional and behavioral outcomes included teacher ratings of emotion regulation, pro-social behavior, and inattention on the Revised Teacher Social Competence scale (R-TSC; Conduct Problems Prevention Research Group, 1995) and the Conners' DSM-IV Inattention scale (Conners, 2001). T-scores accounting for differences in age and gender were utilized in analyses of this latter measure.

Academic outcomes included the *Academic Competence* subscale of the R-TSC scale described above and the Star Early Literacy/Reading and Math computerized assessment (STAR) published by Renaissance Learning. The Academic Competence scale includes teacher-rated items related to setting and achieving goals, solving math problems, reading and answering questions, and turning in homework. The STAR is a nationally-normed computerized adaptive test in which item difficulty is automatically adjusted to reflect students' skill levels, and was designed for repeated administration in order to assess small changes in skill level (i.e. curriculum based measurement). Standard scores used in analyses reflect students' underlying ability level on a continuous vertical scale spanning grade levels, and are a useful measure of absolute growth. Students in kindergarten complete the Early Literacy assessment, while older students are administered Reading, although if they fail screening items they are given the Early Literacy test instead. The Math test is only administered to 1st and 2nd graders. Due to these administration limitations, data are not available for students at all grade levels on both measures.

Findings / Results:

Teacher scores on CLASS indicators of positive behaviors fall in what is considered the moderately high range, with limited room for improvement; scores tapping negative behaviors were also quite low. Initial examination of baseline data (see Table 1) also indicates that students exhibited average or better scores on the R-TSC and Inattention scale, reflecting a normative sample.

Baseline equivalence. There were no differences between intervention and control teachers on any observational scores at pre-test or other classroom-level variables such as teacher experience or class size. Nor were any significant baseline group differences found on student demographics or social-emotional or behavioral measures (although Emotion Regulation approached significance, $p = .08$, with higher scores for the intervention group). Students in the intervention group were rated as having higher baseline Academic Competence and Math scores and lower Early Literacy scores ($p < .05$); however, our statistical models, which focus on the prediction of change over time, account for such baseline differences (which are more likely to be observed in a cluster-randomized, vs. person-randomized, design).

Teacher Outcomes. Results indicate a significant main effect of the intervention on Positive Climate ($b = .77$, $p = .007$), reflecting that teachers in the intervention group were rated as having significantly more positive change in classroom climate than control teachers over time. Contrary to expectation, for the TCI Competent outcome, a significant Experimental Condition × Grade Level interaction effect was observed ($b = -10.78$, $p = .005$), such that second grade teachers in the control group demonstrated significantly greater change in Competence ratings over time than second grade teachers in the intervention group. No significant intervention effects were observed for the other two observational variables.

Student Social-Emotional and Attention Outcomes. Student outcome analyses identified no main intervention effects; however, several statistically significant moderated effects emerged. There was a significant interaction of intervention, grade level, and free/reduced lunch status ($b = -0.42$, $p = .036$), such that a beneficial effect of intervention for the R-TSC Emotion Regulation outcome was observed for kindergarteners receiving free/reduced lunch ($p = .036$) but not for those who do not receive free/reduced lunch ($p = .96$; see Figure 1). A similar pattern of findings was observed for the R-TSC Prosocial Behavior outcome ($b = -0.91$, $p = .005$; Figure 2), such that kindergarteners who received free/reduced lunch significantly benefited from the intervention ($p = .0498$) but those not receiving free/reduced lunch did not significantly benefit ($p = .56$). First graders benefited significantly from the intervention on this outcome regardless of free/reduced lunch status ($p = .033$). No significant intervention effects were observed for the teacher-rated Conners' scale outcome.

Student Academic Outcomes. For the STAR Reading outcome there was a significant Experimental Condition × Grade Level × Time interaction ($b = -119.13$, $p = .011$); in first grade, the intervention group demonstrated significantly greater improvement over time in reading than the control group ($p = .004$), whereas in second grade, change in reading achievement was not significantly different for the intervention and control groups ($p = .67$; see Figure 3). For the STAR Math outcome, when the variance of the random slope at the classroom level was fixed to zero (consistent with the final Reading model), the intervention effect failed to reach significance. However, a random slope model was associated with a significant interaction effect ($b = -46.13$, $p = .049$) in a direction contrary to prediction, such that control students on free/reduced lunch showed greater academic growth over time than intervention students ($p = .02$). No significant intervention effects were observed for the kindergarten Early Literacy outcome or for teacher ratings of Academic Competence.

Discussion and Conclusions:
Overall, this study demonstrates modest effects of a relatively low-cost universal teacher training intervention on classroom climate, student social-emotional outcomes, and reading achievement. It contributes to the efficacy literature for the Incredible Years teacher program, which is supported for use with at-risk children in early childhood settings in combination with other interventions, by showing independent intervention effects in a large, normative sample of early elementary students attending rural schools. It is also the first study of this program to identify specific effects on an academic test, supporting the link between SEL programs and achievement.

Given the high baseline ratings of classroom climate and teacher behavior management skills in this sample, it is perhaps not surprising that expected intervention effects were seen on only one of four indicators of teacher practice change; however, the size of the effect was large and this construct has important implications for student functioning. It is also possible that the CLASS was not sensitive to changes in specific behavior management practices; future analyses will examine teacher self-report data of strategies to explore this possibility. In general, intervention effects for students were moderated by grade level and free/reduced lunch status in the direction predicted. That is, significant benefit was obtained for kindergarteners and 1st graders on social-emotional outcomes and for 1st graders in reading. Thus, changes in classroom climate appeared to translate into positive outcomes in the domains expected, at least for some students.

Two counter-to-hypothesis findings that emerged are more difficult to interpret. There was evidence that 2nd grade control teachers became more competent over time in supporting classroom management and social-emotional skills than did 2nd grade intervention teachers, and that math achievement of control students on free/reduced may have increased more than that of intervention students on free/reduced lunch, an effect that may have been influenced by the 2nd grade teachers. If such findings are replicated, further research is needed to explore any unique teacher factors and/or developmental differences in students that may account for them.

Appendices

Appendix A. References

Bacolod, M. (2007). Who teaches and where they choose to teach: College graduates of the 1990's. *Education Evaluation and Policy Analysis, 29*, 155.

Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (2010). The effects of a multiyear universal social–emotional learning program: The role of student and school characteristics. *Journal of Consulting And Clinical Psychology*, *78*(2), 156-168. doi:10.1037/a0018607

Creemers, B.P.M. (1994). Effective instruction: An empirical basis for a theory of educational effectiveness. In D. Reynolds, B. Creemers, P. Nesselrodt, E. Schaffer, S., Stringfield, & C. Teddlie (Eds). *Advances in School Effectiveness Research and Practice* (pp. 189-206). New York: Elsevier Science Inc.

Department of Education (2008). Education Research CFDA 84.305A-2.

Fixsen, D., Naoom, S., Blase, K., Friedman, R., & Wallace, F. (2005). *Implementation research: A sythesis of the literature*. Tampa, FL: University of South Florida.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal* 38(4), 915 - 945.

Pianta, R. C., K, LaParo, & Hamre, B. (2005). Classroom Assessment Scoring System: University of Virginia.

Pianta, R. C., La Paro, K. M., Payne, C., Cox, M. J., & Bradley, R. (2002). The Relation of Kindergarten Classroom Environment to Teacher, Family, and School Characteristics and Child Outcomes. *Elementary School Journal, 102*(3), 225-238.

Rose, D. J., & Church, R. J. (1998). Learning to teach: The acquisition and maintenance of teaching skills. *Journal of Behavioral Education, 8*, 5–35.

Stringfield, S. (1994). A model of elementary school effects. In D. Reynolds, B. Creemers, P. Nesselrodt, E. Schaffer, S., Stringfield, & C. Teddlie (Eds). *Advances in School Effectiveness Research and Practice* (pp. 153-188). New York: Elsevier Science Inc.

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2001). Preventing conduct problems, promoting social competence: a parent and teacher training partnership in head start. *Journal of Clinical Child Psychology, 30*(3), 283-302.

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology*, *33*(1), 105-124

Webster-Stratton, C., Reid, J., & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: Evaluation of the Incredible Years teacher and child training programs in high-risk schools. *Journal of Child Psychology and Psychiatry*, *49*(5), 471-88.

Appendix B. Tables and Figures

Table 1.
Unadjusted Descriptive Statistics for Teacher and Student Outcome Variables (Means and Standard Deviations)

Measure/Subscale	Group			
	Intervention		Control	
	Pre	Post	Pre	Post
Teacher Outcomes				
CLASS				
Positive Climate	5.46 (.84)	5.82 (.75)	5.63 (.77)	5.54 (.84)
Negative Climate	1.24 (.55)	1.07 (.39)	1.21 (.37)	1.20 (.42)
Behavior Management	5.33 (.89)	5.45 (.87)	5.32 (.92)	5.38 (.87)
TCI				
Competent	4.44 (.50)	4.70 (.50)	4.40 (.68)	4.61 (.60)
Student Outcomes				
R-TSC				
Emotion Regulation	3.64 (.99)	3.79 (.94)	3.54 (.96)	3.64 (.98)
Prosocial Behavior	3.42 (1.13)	3.62 (1.09)	3.23 (1.14)	3.46 (1.15)
Academic Competence	3.36 (1.20)	3.72 (1.15)	3.11 (1.17)	3.66 (1.15)
Conners' Inattention	53.28 (12.63)	51.91 (11.39)	53.57 (12.74)	52.08 (11.22)
STAR Early Literacy	552.40 (113.61)	612.20 (136.74)	570.16 (108.30)	648.04 (107.54)
STAR Reading	211.47 (139.95)	278.06 (241.74)	193.84 (142.43)	248.50 (156.22)
STAR Math	393.47 (120.74)	450.80 (122.13)	362.00 (112.72)	430.92 (122.12)

Note: CLASS scores rated 1-7 with higher being better. TCI Harsh and Competent item average scores range from 0-5, with higher being better for Competent and lower scores being better for TCI Harsh. R-TSC scores range from 1-5, with higher being better. Conners' Inattention scores are T-scores based on age and sex. STAR standard scores reflect ability on a continuous vertical scale scanning grade levels.

Figure 1. Predicted Change in Emotion Regulation by grade and free/reduced lunch status

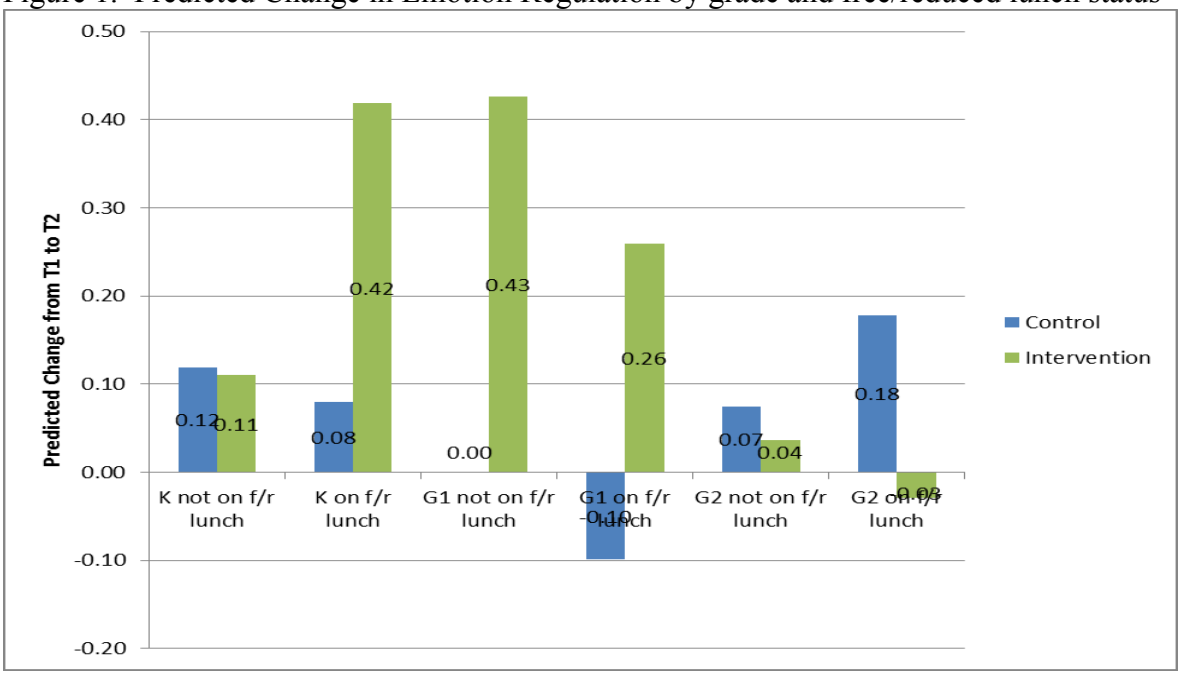

Figure 2. Predicted change in Prosocial Behavior by grade and free/reduced lunch status

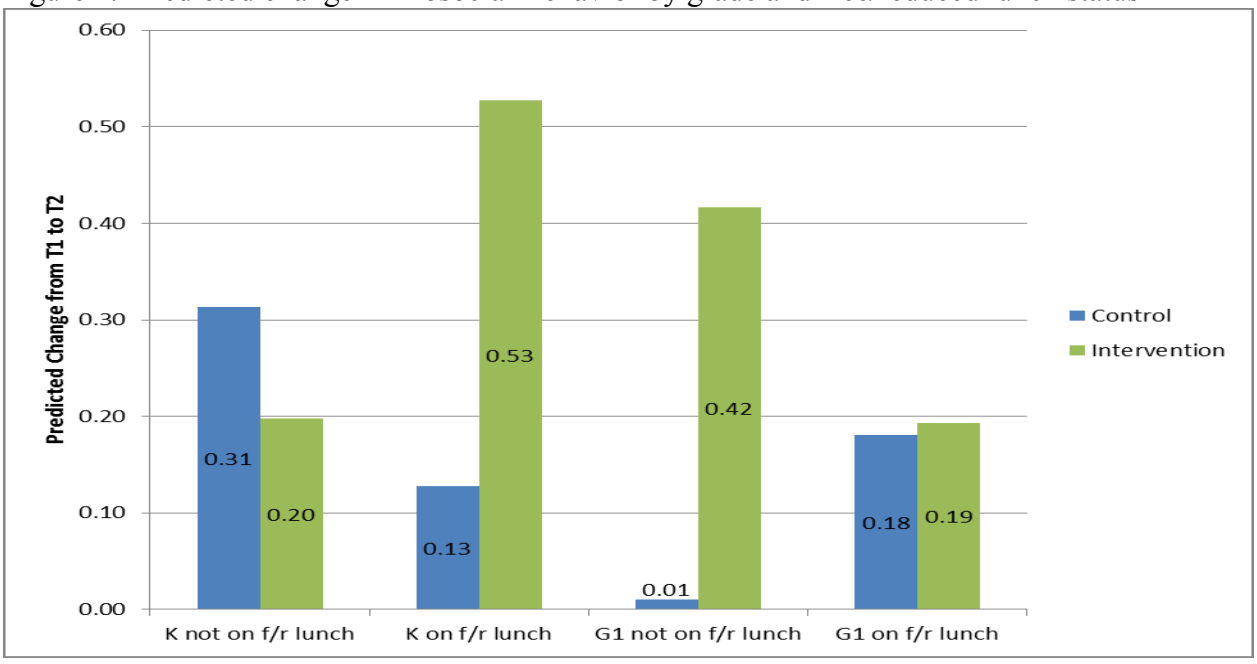

Figure 3. Predicted change in STAR Reading by grade

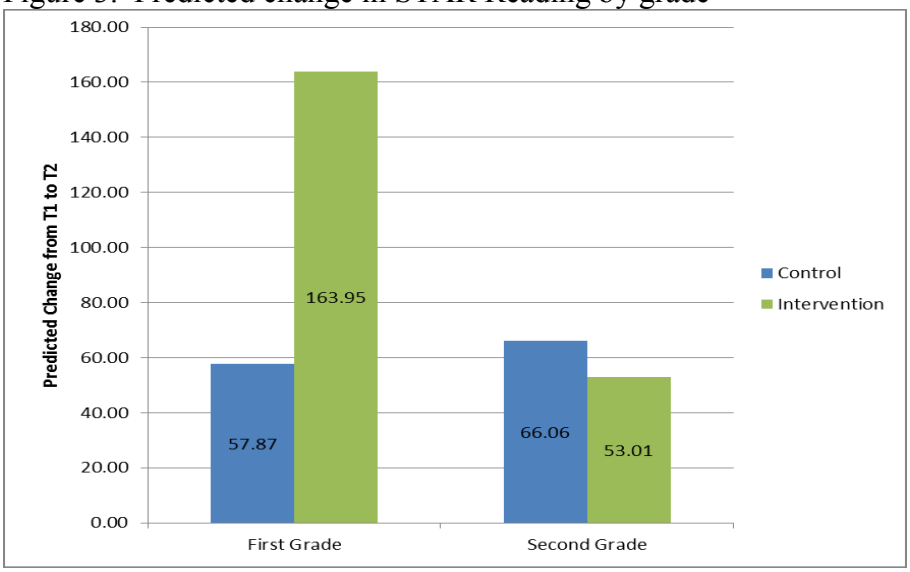

Figure 4. Predicted change in STAR Math by free/reduced lunch