

Contract No.: ED-01-CO-0039
MPR Reference No.: 6031-051

**Social and Character
Development Research
Program National
Evaluation**

***Supporting Statement for
Request for OMB Approval of
SACD Evaluation – Part A***

Original Submission: February 20, 2004

Prepared by:

John Burghardt

Laura Kalb

Larry Snell

Peter Schochet

Susanne James-Burdumy

Revised Submission: December 15, 2006

Prepared by: Amy Silverman

Submitted to:

Teaching and Learning Division
National Center for Education Research
Institute of Education Sciences
U.S. Department of Education
555 New Jersey Ave., NW
Washington, DC 20208

Submitted by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005

Project Officer:

Amy Silverman

Project Director:

John Burghardt

CONTENTS

	Page
PAPERWORK REDUCTION ACT SUBMISSION SUPPORTING STATEMENT.....	1
A. JUSTIFICATION.....	2
1. Circumstances Necessitating Collection Of Information.....	3
2. How, By Whom, And For What Purpose Information Is To Be Used.....	9
3. Use of Automated, Electronic, Mechanical, or Other Technological Collection Techniques.....	19
4. Efforts to Avoid Duplication of Effort.....	19
5. Sensitivity to Burden on Small Entities.....	20
6. Consequences to Federal Program or Policy Activities if the Collection is Not Conducted or is Conducted Less Frequently than Proposed.....	21
7. Special Circumstances.....	21
8. Federal Register Announcement and Consultation.....	21
9. Payment or Gift to Respondents.....	22
10. Confidentiality of the Data.....	23
11. Additional Justification for Sensitive Questions.....	25
12. Estimates of Hour Burden.....	26
13. Estimate of Total Annual Cost Burden to Respondents or Record- Keepers.....	29
14. Estimates of Annualized Cost to the Federal Government.....	30
15. Reasons for Program Changes or Adjustments.....	30
16. Plan for Tabulation and Publication and Schedule for Project.....	30
17. Approval Not to Display the Expiration Date for OMB Approval.....	50
18. Exception to the Certification Statement.....	50

**PAPERWORK REDUCTION ACT SUBMISSION
SUPPORTING STATEMENT**

Agency: U.S. Department of Education, Institute of Education Sciences (IES)

Title: Social and Character Development (SACD) Research Program National Evaluation

Child Report

Teacher Report on Students

Teacher Report on Classroom and School

Primary Caregiver Report

SACD-Activities Observation Instrument

SACD-Activities Principal Interview

School Records Request

A. JUSTIFICATION

This request for OMB clearance addresses data collection activities during the final year of the Social and Character Development (SACD) Research Program that will be occurring in Spring 2007. Original OMB clearance (OMB No.: 1850-0792) encompassed five waves of data collection (Fall 2004, Spring 2005, Fall 2005, Spring 2006, and Spring 2007) in 72 schools. OMB approval to extend data collection activities into additional schools (for a total of 100 schools) was obtained on 8/19/05 (Appendix A3). To date, four waves of data have been completed and the final wave is planned for Spring 2007. Because OMB clearance expires in May 2007, we are requesting an extension through September 2007 because data collection extends into summer months in many sites. All procedures detailed in the original OMB submission and the 7/11/2005 amendment remains the same. Currently, the evaluation of the SACD Research Program is in its third and final year.

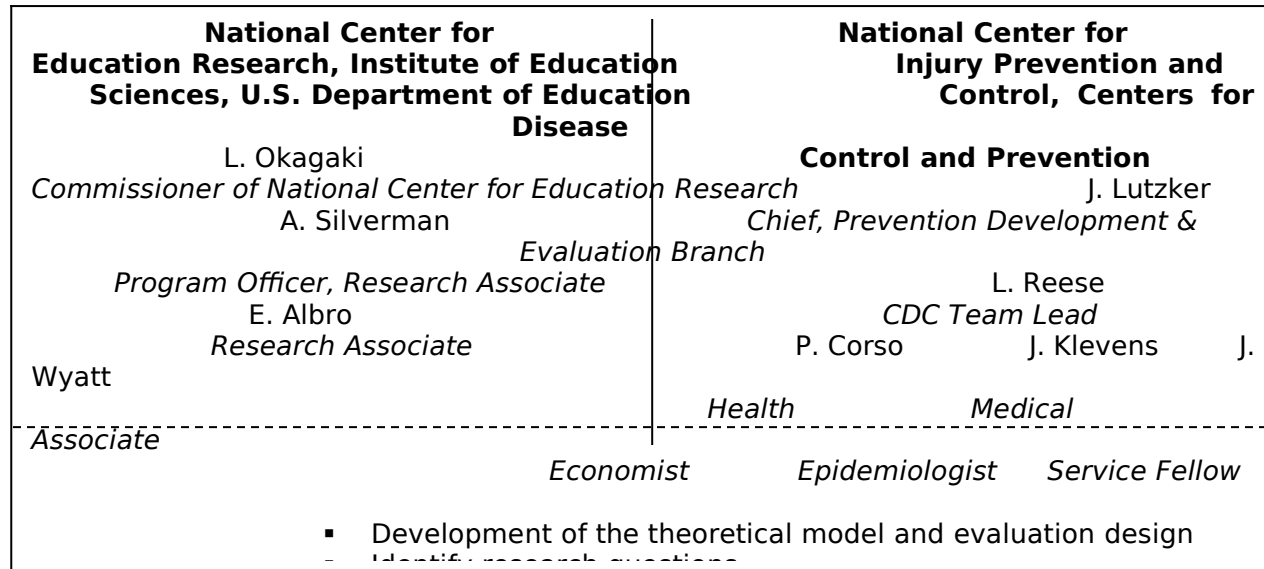
The purpose of the SACD Research Program is to implement and evaluate school-based interventions designed to promote positive social and character development among elementary schoolchildren. Specifically, the program aims to increase positive behaviors, reduce negative behaviors, and ultimately improve children's academic performance. The research will determine, through randomized field trials, whether one or more social and character development program interventions produce meaningful effects among elementary schoolchildren. During FY2003, the Institute of Education Sciences, U.S. Department of Education funded seven grantees to examine SACD intervention programs across eight sites. Grantees under the SACD program are responsible for implementing one or more identified SACD programs and working with the national evaluator to facilitate collecting data at each site in the fall and spring of the third grade year, in the fall and spring of the fourth grade year, and the spring of the fifth grade year (see Figure 1 for an organizational chart of the program).

The *No Child Left Behind (NCLB) Act* of 2001, Pub. L. No. 107-110, 115 STAT.1425, enacts the Partnerships in Character Education program (administered through the Department's Office of Safe and Drug Free Schools) to support the design and implementation of instruction directed toward promoting aspects of character (such as citizenship, respect, and responsibility) to, in turn, improve the school environment. The legislation requires that education decision-makers base instructional practices and programs on scientifically based research. Such research has been limited, however; particularly, evidence from rigorous evaluations utilizing randomized experimental designs. In response to the need for rigorous evaluations of school-based programs that promote positive character development and reduce school violence and other antisocial behaviors, the U.S. Department of Education's Institute of Education Sciences is supporting the SACD Program, in collaboration with the National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

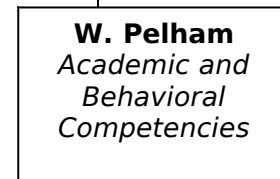
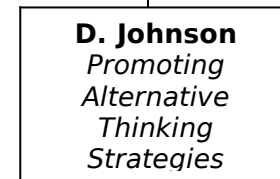
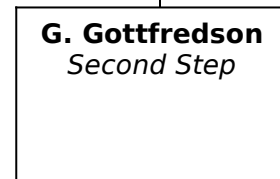
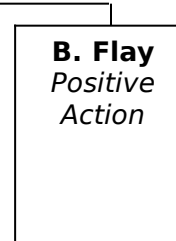
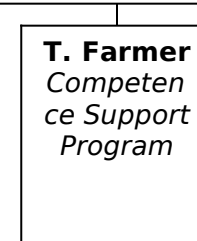
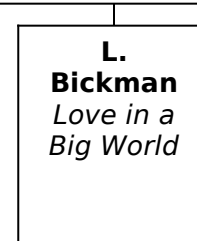
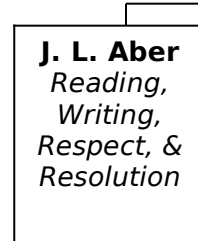
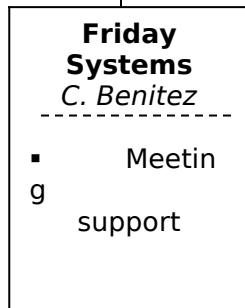
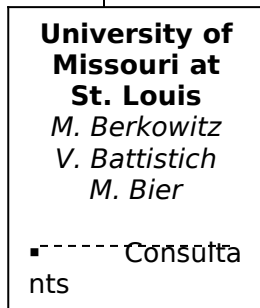
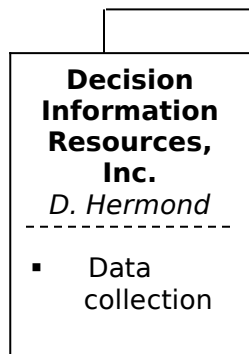
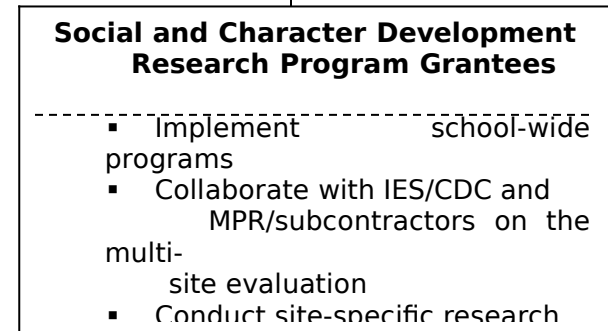
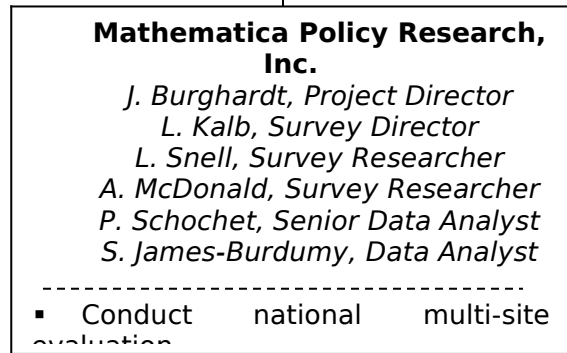
1. Circumstances Necessitating Collection of Information

During the past decade, an increasing number of school-based initiatives have been implemented to support positive social and character development, promote positive behaviors, prevent negative behaviors, and, ultimately, improve academic achievement. Fourteen states mandate character education and another fourteen have enacted legislation encouraging character education. In addition, 47 states and the District of Columbia had received Federal Character Education Partnership Grants as of 2002.

Figure 1. Social and Character Development Research Program



4



While many of these social and character development initiatives have shown promise, few rigorous evaluations of these school-based interventions have been conducted. Very little scientific evidence currently exists to enable administrators to identify effective programs. By subjecting the most promising interventions to scientific study, the SACD Research Program will make a significant contribution to knowledge of effective practices in the social and character development field. The SACD Research Program will also influence decisions that school administrators make about which interventions to adopt.

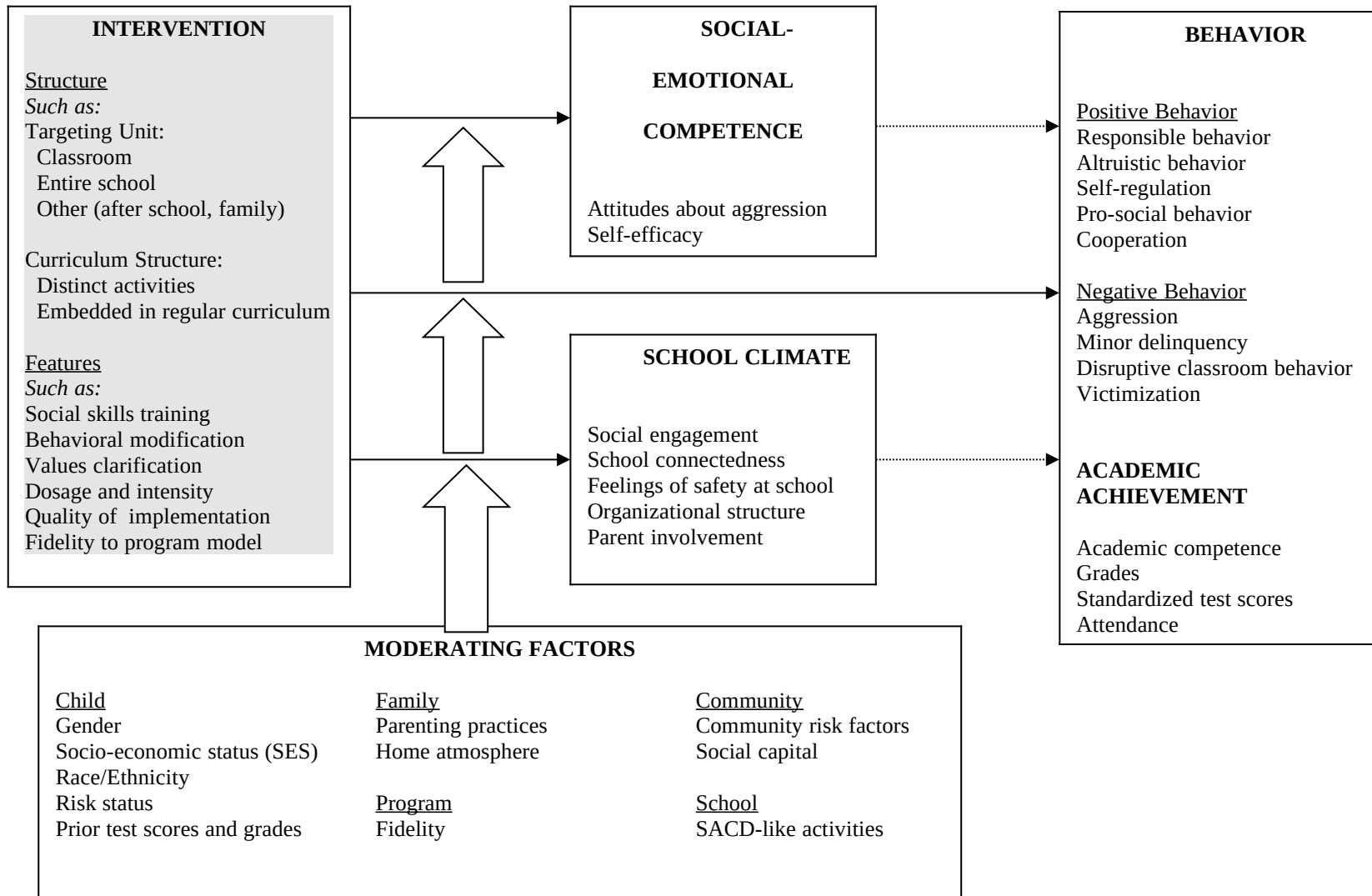
To meet IES/CDC's purpose of identifying effective strategies for enhancing elementary schoolchildren's social and character development, the study provides evidence of the impacts the interventions have on the children they serve relative to the educational experiences prevailing in their communities. At each site, grantees randomly assigned schools to two groups: (1) a treatment group in which the intervention was implemented, and (2) a control group that did not receive the experimental intervention. Children's progress, changes in school climate, and activities to promote social and character development is being evaluated longitudinally through child reports, teacher reports on children in the study, teacher background surveys, other school staff surveys, principal interviews, primary caregiver reports, school records, and school observations. Baseline data was collected in fall 2004. Impacts (treatment-control differences) are being assessed during (1) the spring semester of third grade (spring 2005), (2) the fall semester of fourth grade (fall 2005), (3) the spring semester of fourth grade (spring 2006), and (4) the spring semester of fifth grade (spring 2007). Because the interventions are school-wide, and to account for student mobility, any *new* students who move into the treatment and control group schools during the follow-up period who are in the same grades as the original sample members are considered new entrants and are utilized in outcome analyses.

A variety of child, family, and community characteristics are expected to moderate the effects of the interventions. The measures used in the multisite research program are designed to detect these effects. The multisite analyses will examine intervention elements that directly affect children’s behavior and achievement, as well as the demographic factors that interact with intervention approaches and may also influence how the interventions have their effects. The analyses will also examine the outcomes that may serve as mediators of the ultimate child outcomes, specifically, the social-emotional competence and school climate outcomes indicated in Exhibit 1. The study’s three primary research questions will guide the multisite analysis:

1. ***What are the overall impacts of the SACD initiatives on student- and school-level outcomes across the seven programs combined?*** Which particular outcomes are most affected? How do impacts on students’ attitudes, behaviors, and academic achievement change over time?
2. ***What works, for whom, and under what conditions?*** To what extent do impacts vary across subgroups defined by key structural elements and features of the interventions? To what extent do impacts vary across the interventions being implemented in each site? To what extent do impacts vary across subgroups defined by key student characteristics? Are impacts larger for those who receive a higher “dose” of the treatment than for those who receive a lower dose?

EXHIBIT 1

CONCEPTUAL MODEL FOR THE SACD EVALUATION



3. ***Are impacts on mediating outcomes consistent with impacts on longer-term student outcomes?*** What is the process by which the interventions influence students' behavior and academic achievement?

The first primary research question pertains to the SACD programs' overall impacts. Although the interventions differ across sites, it is of policy importance to examine the overall effectiveness of the social and character development initiatives funded under the SACD Research Program and to examine the particular outcomes the interventions are most likely to affect.

It is important to go beyond the overall impacts to examine what works, for whom, and under what conditions. Thus, the analyses will also examine the impacts of subgroups of similar interventions and impacts of the intervention programs on key subgroups of children. These analyses can provide important information to help improve those interventions and guide their expansion and development, as well as important information on whether and how to target the programs.

The third primary research question focuses on understanding the processes by which the interventions achieve positive outcomes. This information can help program staff focus improvement efforts on program features that are most effective. It is especially important to determine which program features are highly correlated with longer-term, sustainable, positive outcomes.

To address these overall questions, specific data collection activities during each wave include the administration of surveys to children, teachers, principals, and primary caregivers; school observations; and school record abstraction. Exhibit 2 provides a listing of each instrument and the key constructs and measures it includes. Appendix I provides normative and psychometric information on the instruments that will be used in the SACD data collection.

Appendixes II-VIII include copies of the instruments. Appendix IX provides information on which measures require the permission of the original authors for use by others. All Appendixes are located in a separate document.

2. How, by Whom, and for What Purpose Information is to be Used

The purpose of the national multisite evaluation is to determine the overall efficacy of the SACD interventions that the seven grantees are implementing in schools. The multisite evaluation will provide important information to determine which interventions lead to improvements in child outcomes. Specifically, the evaluation will determine which interventions support positive social and character development, promote positive behaviors, reduce negative behaviors, and, ultimately, improve academic achievement. Additionally, the multisite evaluation will identify specific program features that are linked to these impacts and assess under what conditions, and for which children, the interventions are effective. Results from the multisite evaluation will provide school districts and education professionals with the information they need to make informed choices about which social and character development interventions to adopt. The results also will offer policymakers rigorous evidence for use in making decisions about program funding. The results of the multisite evaluation will be presented in annual reports and briefings for policymakers beginning in March 2006. To date, no annual reports have been published. In addition, IES, CDC, and researchers will present the findings at professional conferences. To date, symposium presentations have been given at annual meetings of the Society for Prevention Research (2006), Society for Research in Child Development (2005), and American Evaluation Association (2004).

Site-specific data collected for the multisite evaluation will be provided to the seven grantees following each wave of data collection for use in their site-specific analyses. To date,

each grantee site has applied for and received a restricted-use data license for their site-specific data from the National Center for Education Statistics (NCES) who is responsible for reviewing and approving the licenses, monitoring data security, and protecting the confidentiality of the datasets. For more information related to the policies and procedures involved in obtaining restricted-use licenses, please refer to the following website: <http://nces.ed.gov/StatProg/confproc.asp>. IES does not have copies of the signed restricted-use licenses on file because these are administered and monitored solely by NCES (contact person is Marilyn Seastrom, Chief Statistician and Program Director).

Data collected for the multisite evaluation will also be made available to researchers for secondary analyses on a restricted basis after the multisite evaluation analyses have been conducted and reported. The datasets will be made available to qualified researchers who agree to follow specified practices for ensuring confidentiality. NCES will be responsible for approving and issuing restricted-use data licenses to qualified researchers utilizing a set of standardized procedures and policies developed specifically for these purposes. For more information on these policies and procedures, please refer to the NCES website in the previous paragraph. The data to be collected are described in the following paragraphs.

To evaluate the effectiveness of the selected SACD programs, the national evaluator (Mathematica Policy Research, Inc., or “MPR”) has been coordinating data collection from each of the eight research sites (six of the grantees have a single research site; one grantee has two research sites). Because the data collected at each site is being combined and compared with the data collected from other sites, data collection procedures must be uniform across all of the sites. Joint meetings of IES, CDC, the national contractor, and grantee staff have been held twice a year to facilitate the development of data collection protocols that ensure consistency in

procedures while meeting the needs of both the grantees' site-specific work and the national contractor's multisite evaluation responsibilities. Biweekly conference calls between meetings also provide a forum for making adjustments to the protocols if needed.

The measures presented in Exhibit 2 capture key aspects of the theoretical model presented in Exhibit 1. All measures have been administered uniformly at all grantee sites. The data collected from children, teachers, primary caregivers, school staff, principals, school records, and school observations will be used to:

- Obtain data on classroom and teacher characteristics and program features that are not otherwise available and that are necessary to analyze implementation of the SACD programs;
- Obtain data necessary to interpret findings with respect to the impact of the various SACD interventions across all grantee research sites;
- Obtain outcome data on children's behavior that are not otherwise available and are necessary to analyze the impacts of the social and character development programs across all grantee research sites.

A brief description of each data instrument that will be administered during 2006-2007 data collections provided below. The appendices are found in a separate document.

Child Report (Appendix II). The child report will be administered to groups of 15 to 20 children at a time. It is estimated to take 50 minutes, including time to distribute and collect the report booklets and provide instructions. Assessors have been trained to administer the child reports uniformly across each research site. The child report is administered during all five waves of data collection, with fall 2004 serving as the baseline.

Teacher Report on Students (Appendix III). This is a paper and pencil rating which will be completed by the teacher of each child's social and academic competence, conduct, and behavior—all of which are key outcomes for analysis. During spring 2006-2007 data collection, the Teacher Report is estimated to take 16 minutes per child to complete. These reports are collected by the grantees during all five data collection periods.

EXHIBIT 2
COMPONENTS AND SOURCES OF DATA COLLECTION INSTRUMENTS

Instrument/ Time	Components	Broad Construct	Source
Child Report <ul style="list-style-type: none"> • Fall 2004 • Spring 2005 • Fall 2005 • Spring 2006 • Spring 2007 	Normative Beliefs About Aggression	Attitudes about aggression	Huesmann, L.R., & Guerra, N.G. (1997). Children's normative beliefs about aggression and aggressive behavior. <i>Journal of Personality and Social Psychology</i> , 72, 408-419.
	Children's Self-Efficacy for Peer Interaction Scale	Self-efficacy	Wheeler, V. A., & Ladd, G. W. (1982). Assessment of children's self-efficacy for social interactions with peers. <i>Developmental Psychology</i> , 18, 795-805.
	Children's Empathy Questionnaire	Empathy	Funk, J., Elliott, R., Bechtoldt, H., Pasold, T., & Tsavoussis, A. (2003). The Attitudes Toward Violence Scale: Child version. <i>Journal of Interpersonal Violence</i> , 18, 186-196.
	Engagement versus Disaffection with Learning	School engagement	Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. <i>Journal of Educational Psychology</i> , 95, 148-162.
	Sense of School as a Community Scale; Child Version	School connectedness	Roberts, W., Horn, A., & Battistich, V. (1995, April). Assessing students' and teachers' sense of the school as a caring community. Paper presentation at the meeting of the American Educational Research Association.
	Self-Report of Prosocial Behavior	Child's prosocial behavior	Soloman, D., Battistich, V., Watson, M. Schaps, E., & Lewis, C. (2000). A six-district study of educational change: Direct and mediating effects of the Child Development Project. <i>Social Psychology of Education</i> , 4, 3-51.
	Feelings of safety at school	Feelings of safety at school	Items provided by IES/CDC.
	Aggression Scale	Child's aggressive behavior	Orpinas, P., & Frankowski, R. (2001). The Aggression Scale: A self-report measure of aggressive behavior for young adolescents. <i>Journal of Early Adolescence</i> , 21, 50-67.
	Frequency of Delinquent Behavior	Minor Delinquency	Adapted from: Loeber, R., & Dishion, T.J. (1983). Early predictors of male delinquency: A review. <i>Psychological Bulletin</i> , 94, 325-382
	Victimization	Victimization in school	Orpinas, P., & Kelder, S. (1995). Students for Peace Project: Second student evaluation. Unpublished Manuscript. Houston, TX: University of Texas Health Science Center at

Instrument/ Time	Components	Broad Construct	Source
			Houston, School of Public Health.
Teacher Report on Sample Children	Social Competence	Child's self- regulation, cooperation, and prosocial behavior	Conduct Problems Prevention Research Group (1999). Initial impact of the Fast Track prevention trial for conduct problems I: The high-risk sample. <i>Journal of Consulting and Clinical Psychology</i> , 67, 631-647.
<ul style="list-style-type: none"> • Fall 2004 • Spring 2005 • Fall 2005 • Spring 2006 • Spring 2007 	Responsibility Scale; Teacher Report	Child's responsibility	Items developed by IES/CDC.
	Parent and Teacher Involvement Measure; Teacher Report	Parent involvement in the child's school life	CPPRG (1991). <i>Parent-Teacher Involvement Measure - Parent</i> . (Online). Available: http://www.fasttrackproject.org/
	Report of Prosocial Behavior	Child's prosocial behavior	Soloman, D., Battistich, V., Watson, M. Schaps, E., & Lewis, C. (2000). A six-district study of educational change: Direct and mediating effects of the Child Development Project. <i>Social Psychology of Education</i> , 4, 3-51.
	BASC Aggression Subscale; Teacher Report	Child's aggressive behavior	Reynolds, C.R., & Kamphaus, R.W. (1998). <i>Behavioral Assessment System for Children</i> . Circle Pines, MN: American Guidance Service Inc.
	BASC Conduct Problems Subscale; Teacher Report	Child's conduct problems	Reynolds, C.R., & Kamphaus, R.W. (1998). <i>Behavioral Assessment System for Children</i> . Circle Pines, MN: American Guidance Service Inc.
	Sutter-Eyberg Student Behavior Inventory	Disruptive classroom behavior	Funderburk, B.W., & Eyberg, S.M. (1989). Psychometric characteristics of the Sutter-Eyberg Student Behavior Inventory: A school behavior rating scale for use with preschool children. <i>Behavioral Assessment</i> , 11, 297-313.
	SSRS Academic Competence and Achenbach Teacher Report Form (TRF)	Academic competence	Adapted from: Gresham, F.M., & Elliott, S.N. (1990). <i>Social Skills Rating System</i> . Circle Pines, MN: American Guidance Service. Achenbach, T. M. (1991). Manual for the teacher's report form and 1991 profile. Burlington, VT: University of Vermont, Department of Psychiatry.
Teacher Background Survey	Teacher Survey on Professional Development and Training	Demographics , teaching background, type of	Lewis, L. et al. (1999). U.S. Department of Education. National Center for Education Statistics. <i>Teacher Quality: A report on the</i>

Instrument/ Time	Components	Broad Construct	Source
<ul style="list-style-type: none"> • Fall 2004 • Spring 2005 • Fall 2005 • Spring 2006 • Spring 2007 		certification, professional development activities	<i>Preparation and Qualifications of Public School Teachers</i> . Washington, DC: NCES 1999-080.
Other School Staff Survey	School-Level Environment Questionnaire	School Organizational Climate	Rentoul, A.J., & Fraser, B.J. (1983). Development of a school-level environment questionnaire. <i>Journal of Educational Administration</i> , 21, 21-39.
<ul style="list-style-type: none"> • Fall 2004 • Spring 2005 • Fall 2005 • Spring 2006 • Spring 2007 	Feelings of Safety at School	Feelings of safety at school	Fisher, D. L., & Fraser, B. J. (1991). Validity and use of school environment instruments. <i>Journal of Classroom Interaction</i> , 26, 13-18. Items provided by IES/CDC.
	Teacher Survey on Professional Development and Training	Demographics, teaching background, type of certification, professional development activities	Lewis, L. et al. (1999). U.S. Department of Education. National Center for Education Statistics. <i>Teacher Quality: A report on the Preparation and Qualifications of Public School Teachers</i> . Washington, DC: NCES 1999-080.
Primary Caregiver Report	BASC Aggression Subscale; Parent Report	Child's aggressive behavior	Reynolds, C.R., & Kamphaus, R.W. (1998). <i>Behavioral Assessment System for Children</i> . Circle Pines, MN: American Guidance Service Inc.
<ul style="list-style-type: none"> • Fall 2004 • Spring 2005 • Fall 2005 • Spring 2006 • Spring 2007 	BASC Conduct Problems Subscale; Parent Report	Child's conduct problems	
	Community Risks	Community risk	Forehand, R., Brody, G.H., Armistead, L. et al. (2000). The role of community risks and resources in the psychosocial adjustment of at-risk children: An examination across two community contexts and two informants. <i>Behavior Therapy</i> , 13, 395-414.
	Confusion, Hubbub, and Order Scale	Environmental confusion	Matheny, A.P., Wachs, T.D., Ludwig, J.L., & Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the Confusion, Hubbub, and Order Scale. <i>Journal of</i>

Instrument/ Time	Components	Broad Construct	Source
			<i>Applied Developmental Psychology</i> , 16, 429-444.
	Alabama Parenting Questionnaire	Positive parenting and supervision/ monitoring	Shelton, K.K., Frick, P.J., & Wootton, J. (1996). Assessment of parenting practices in families of elementary school-age children. <i>Journal of Clinical Child Psychology</i> , 25, 317- 329.
	Report of Prosocial Behavior	Child's prosocial behavior	Soloman, D., Battistich, V., Watson, M. Schaps, E., & Lewis, C. (2000). A six-district study of educational change: Direct and mediating effects of the Child Development Project. <i>Social Psychology of Education</i> , 4, 3-51.
	Child-Centered Social Control	Social capital in the community	Sampson, R.J., Morenoff, J.D., & Earls, F. (1999). Beyond social capital: Spatial dynamics of collective efficacy for children. <i>American Sociological Review</i> , 64, 633-660.
	Social Competence	Child's self- regulation, cooperation, and prosocial behavior	Conduct Problems Prevention Research Group (1999). Initial impact of the Fast Track prevention trial for conduct problems I: The high-risk sample. <i>Journal of Consulting and Clinical Psychology</i> , 67, 631-647.
	Responsibility Scale; Parent Report	Child's responsibility	Items developed by IES/CDC.
	Parent and Teacher Involvement Measure; Parent Report	Parent involvement in the child's school life	CPPRG (1991). <i>Parent-Teacher Involvement Measure - Parent</i> . (Online). Available: http://www.fasttrackproject.org/
	Background Questionnaire	Demographics	CDC

Teacher Report on Classroom and School (Appendix IV). Questions from three teacher surveys in the original protocol (the Teacher Background Survey, Other School Staff Survey, and SACD-Activities Teacher Survey) have been combined into one survey for 2005-2006 and 2006-2007 data collection (the Teacher Report on Classroom and School). This change is described in the 7/11/05 Burden Hour Memorandum (Appendix A2). All third, fourth, and fifth grade teachers in study schools will complete a self-administered questionnaire on the

organizational climate of the school, the social and character development-like activities implemented in their classroom, and their own professional background. The Teacher Report on Classroom and School is estimated to take approximately 33 minutes to complete. Teachers complete the report while their students are filling out the child report. Part I of the survey (previous Teacher Background Survey) covers basic demographic characteristics (gender, race, and ethnicity), experience in the field of education, type of certification, educational background, and professional development activities. Grantees are responsible for distributing and collecting the reports from teachers of children in the sample.

Obtaining teacher background information is critical for several reasons. First, in order to characterize the study sample, the national evaluator needs to be able to describe the teachers who participated in the study and provided reports on children in the sample. Second, this information is important to determine whether there are any moderating effects of teacher background on children's outcomes.

Part II of the Teacher Report on Classroom and School (Previous School Staff Report) gathers basic demographic data from respondents, as well as ask them about their views on connectedness within the school, the school's organizational climate, and school safety. It also requests information about their recent professional development activities. The Teacher Report on Classroom and School will be distributed and collected by the grantees during each of the five data collection periods.

Part III of the Teacher Report on Classroom and School consists of the SACD-Activities Teacher Survey. This survey complements the school observation data and principal interview. It will be a written survey that includes questions about school activities that may influence the social and character development of the students (e.g., use of official character education

curricula and conflict resolution activities). The report will be administered during all five waves of data collection.

Primary Caregiver Report (Appendix V). The primary caregiver report will gather important information from the adult caregiver primarily responsible for each child in the sample. This self-administered questionnaire will take an estimated 20 minutes to complete. Topics covered include demographic information about the caregiver, information on the family and neighborhood environment, information about the child's behavior, attitudes toward parenting, and involvement in the school activities of the child. The primary caregiver report will be administered during all five waves of data collection in order to track any changes in children's home or community environment, their behavior at home, and the caregivers' involvement in their children's lives. The grantees will arrange to have the teachers of children in the sample distribute and collect sealed envelopes containing the primary caregiver report. Computer-assisted telephone interviews (CATI) will be conducted by the national evaluator with survey non-respondents in order to maximize the response rate.

SACD-Activities Observation Instrument (Appendix VI). The SACD-Activities Observation Instrument includes a variety of measures designed to document SACD activities and strategies occurring in each school. Each school will be observed by a member of the national evaluator's staff. The observation protocols designed to gather information on the school climate, such as cleanliness, graffiti, evidence of disruption, violence, and misbehavior. It is also designed to gather information on the type of social and character development activities occurring at the school (e.g., posters on the wall promoting positive character traits, awards hung for good classroom behavior). It will be administered during all five waves of data collection. The observation will not involve any school staff time and thus is not included in

estimates of burden provided below. Observations of classrooms will take place when no children are present.

SACD-Activities Principal Interview (Appendix VII). The interview with principals will complement the school observation data and ask about related topics. The interview will be conducted in a semi-structured format and include both closed- and open-ended questions. MPR staff will conduct the interview over the phone with principals in order to gather information about school activities that may influence the social and character development of the students (e.g., use of official character education curricula and conflict resolution activities). The interview with principals is estimated to take 45 minutes and will be administered during all five waves of data collection.

School Records Request (Appendix VIII). The school records data will complement the school observation and principal interview data and will provide important information on child outcomes. The purpose of obtaining school-level information from records about enrollment numbers and characteristics, staffing, children receiving specific services, and behavior problems is to help characterize the basic aspects of the school environment for staff and students, and understand how the school environment changes as a function of a social and character development intervention implementation. The child-level school records also provide important academic and behavioral outcome variables that are expected to change after implementation of an intervention.

Because school records vary among school districts and schools, we will first determine what records are available for each school from the list of records in Appendix X. After determining which records are available for all schools, grantees will collect those common records. They will send this data to the national evaluator electronically. If key items are missing, data collectors will be sent to manually collect the missing records. We assume that the

burden on school and district staff of providing electronic files with the requested data or of providing access to paper files for field staff to extract data will average four hours per school.

3. Use of Automated, Electronic, Mechanical, or Other Technological Collection Techniques

The data collection plan reflects sensitivity to issues of efficiency, accuracy, and respondent burden. Where feasible, information will be gathered from existing data sources, such as school records, using straightforward reporting forms. School records information will be gathered via computer files if a school prefers this method. However, most data can be obtained only from students, their caregivers, their teachers, and the school staff who work with them.

Technological tools will be used to minimize respondent burden whenever possible. The phone number and electronic mail address for the national evaluator will be included on the front of each self-administered questionnaire in order for respondents to easily advance any queries they might have. A computer-assisted telephone interviewing (CATI) system will be used for interviews with non-respondents to the self-administered primary caregiver report. This will increase the efficiency of the interview and thus reduce the time required of those who did not initially complete the report.

4. Efforts to Avoid Duplication of Effort

A literature search on school-based interventions designed to promote positive behavior and/or reduce negative behavior identified various studies that examined how students tend to respond to these interventions. However, the vast majority of studies included pretest-posttest designs or quasi-experimental methods. There is a significant lack of studies that systematically and rigorously evaluate the impact of social and character development intervention programs utilizing randomized field trials—the preferred methodology for answering causal questions about the effectiveness of programs.

In addition, a unique feature of this program is that a core set of measures will be used to collect consistent data across multiple social and character development interventions. One weakness of previous research programs has been that primary outcomes of interest, such as social competencies, aggression, and prosocial behavior, have been measured inconsistently across programs. Thus, when reviewing research findings from multiple evaluation studies, it is difficult to come to a consensus about “what works” to prevent or promote specific competencies and behaviors. By using a core set of measures in the current study, we will be able to examine and compare the effectiveness of a variety of interventions on the same student competencies and behaviors.

IES has communicated frequently on this and similar projects related to social and character development school-based interventions with experts in the field. In these communications, although there has been much interest in research of this kind, no similar efforts have been identified.

5. Sensitivity to Burden on Small Entities

The primary entities for the study are schools and the children who attend them. Burden is reduced for all respondents by requesting only the minimum information required to meet the study objectives. The burden on schools has been minimized through the careful specification of information needs, restricting questions to generally available information, and the design of the data collection strategy, particularly the survey methods, to minimize burden on respondents. All multisite data collection has been coordinated by the evaluation contractor, Mathematica Policy Research, Inc. (MPR), and its subcontractor, Decision Information Resources, Inc. (DIR), so as to minimize the burden on school staff, children, and their primary caregivers.

6. Consequences to Federal Program or Policy Activities if the Collection is Not Conducted or is Conducted Less Frequently than Proposed

Without the data from the national evaluation, IES/CDC will be unable to assess the impacts of specific school-based interventions on social and character development outcomes. In particular, IES/CDC will not know whether any of the programs had any impacts; nor will they know whether such programs can achieve desired outcomes for students, their caregivers, and their schools more generally. As noted previously, NCLB supports the design and implementation of instruction directed toward promoting aspects of character. But the legislation also requires that education decision-makers base instructional practices and programs on scientifically based research. Without data from the national evaluation, federal resources would have to be allocated and program decisions made in the absence of valid evidence on the effectiveness of the programs. In addition, there is a need for the data to be collected over a three-year time period because school administrators need to know how long interventions need to be implemented to detect increasingly positive outcomes in their students.

7. Special Circumstances

There are no special circumstances involved with this data collection.

8. Federal Register Announcement and Consultation

a. Federal Register Announcement

The 60 day Federal Register notice was published on October 16, 2006 on page 60700. No comments were received during the comment period.

b. Consultations Outside the Agency

A consortium of the SACD grantees has engaged in a review of the overall study design, the data collection plan, and the data collection instruments. They represent a number of the nation's leading researchers in the area of social and character development, as well as national experts on school-based data collection. The consortium includes:

- J. Lawrence Aber, New York University
- Leonard Bickman, Vanderbilt University
- Thomas W. Farmer, University of North Carolina, Chapel Hill
- Brian Flay, University of Illinois at Chicago
- Gary D. Gottfredson, University of Maryland
- Deborah B. Johnson, Children's Institute
- William E. Pelham, Jr., State University of New York, Buffalo

c. Unresolved Issues

None.

9. Payment or Gift to Respondents

Primary caregivers will be providing information unavailable from other data sources, as well as information on child behavior that can be triangulated with that provided by the children and their teachers. Incentives for the 2006-2007 data collection period will be identical to what was described in the initial clearance package. We will offer caregivers \$10 each time a report is completed to compensate for the time and effort dedicated to completing the 15-minute report.

Teachers will also be completing reports on individual children in the sample that are estimated to take approximately 16 minutes each. Teachers will be providing information not being obtained elsewhere, as well as information that can be triangulated with that provided by

the children and their primary caregivers. We will offer teachers \$5 for each child report (or the wage required by their union for such activities) for their time and effort.

Teachers of non-sampled children will be offered \$10 (or the wage required by their union for such activities) for completing the Teacher Report on Classrooms and School. It is estimated to take 33 minutes to complete. Grantees also will offer compensation for teacher's time and effort of participating in the study; no additional compensation is planned for the SACD-Activities instruments.

10. Confidentiality of the Data

All data collection activities will be conducted in full compliance with Department of Education regulations to maintain the confidentiality of data obtained on private persons and to protect the rights and welfare of human research subjects as contained in Department of Education regulations. Research participants (primary caregivers, students, teachers, and school staff) will sign written consent (or assent) forms. The consent materials will inform respondents about the nature of the information that will be requested and confidentiality protection, and they will be assured that information will be reported only in aggregate, statistical form. The consent materials will also inform respondents that the data will be used only for research purposes by researchers who have signed a confidentiality agreement.

Currently, many grantees are in the process of obtaining updated IRB approval for the consent forms they will be using for data collection in Spring 2007. For this reason, we are not able to provide the most up-to-date version of the consent forms. A previous version of the consent forms (or the current ones that I have on record) can be furnished upon request.

At the time of the initial OMB submission in 2004, it was felt that it a System of Records Notice was not necessary to submit. Although the collection asks for the date of birth of the student (year, month, day), it does not ask for name or social security of the child. For this reason, it was felt that a SORN was not necessary to submit due to the other measures of protection that were taken to ensure confidentiality.

In addition to the consent forms, each self-administered instrument will include a reminder on the protection of confidentiality. Where data are collected through in-person interviews or group surveys—for instance, the school principal interview and the child report—interviewers will remind respondents of the confidentiality protections provided, as well as their right to refuse to answer questions to which they object. During the group administration of the child report, desks will be arranged in classroom style to ensure that children cannot see the responses provided by classmates. All data collectors and interviewers will be knowledgeable about confidentiality procedures and will be prepared to describe them in full detail, if necessary, or to answer any related questions raised by respondents.

The national evaluator has a long history of protecting confidentiality and privacy of records, and considers such practice a critical aspect of the scientific and legal integrity of any survey. The integrity the national evaluator brings to protecting data confidentiality and privacy will extend to every aspect of survey operations and data handling in the field for the SACD program. The national evaluator plans to use its ongoing, long-standing techniques that have proven effective in the past. Every interviewer will be required to sign a pledge to protect the confidentiality of respondent data. The pledge indicates that any violation or unauthorized disclosure may result in legal action or other sanctions by the national evaluator. The national evaluator will require that grantees have similar protections in place for the data that they will

assist in collecting. The national evaluator requires all interviewers to view a videotape about the Belmont Report for the protection of human subjects, and includes a discussion of human subject protection as part of their training. After participating in this training, interviewers sign a form certifying that they have received the training. A copy of both pledges will be kept on file and will be available upon request.

In addition, the following safeguards are routinely employed by MPR to carry out confidentiality assurances:

- Access to sample selection data is limited to those who have direct responsibility for providing the sample. At the conclusion of the research, these data are destroyed.
- Identifying information is maintained on separate forms which are linked to the interviews only by a sample identification number. These forms are separated from the interviews as soon as possible.
- Access to the file linking sample identification numbers with the respondents' identification and contact information is limited to a small number of individuals who have a need to know this information.
- Access to the hard copy documents is strictly limited. Documents are stored in locked files and cabinets. Discarded material is shredded.
- Computer data files are protected with passwords and access is limited to specific users. With especially sensitive data, the data are maintained on removable storage devices that are kept physically secure when not in use.

11. Additional Justification for Sensitive Questions

It is not possible to avoid sensitive questions in a study of programs designed specifically to address engagement in disruptive or aggressive behavior. Thus, some questions of a sensitive nature are included in the child and primary caregiver reports. Although these questions may be regarded as sensitive, they are all derived from instruments designed to assess children's

aggressive and prosocial attitudes/behaviors that have been administered to and validated with samples of children and their parents or guardians in past research.

For the children, the most sensitive questions relate to minor delinquent behavior (e.g., cheating)—activities that are likely to be influenced by the intervention. The collection of this information is critical to evaluate the effectiveness of these interventions.

Each group administration will be moderated by at least two trained evaluator staff members. They will have been carefully trained on how to read the questions aloud to the group in a neutral fashion, and how to answer questions from children appropriately. In addition, steps will be taken to ensure that all children can answer questions about their behavior confidentially, and that their names do not appear on the completed instruments. Finally, during the assent procedure at the beginning of the group administration, we will explain to the children that they can skip any questions that they feel uncomfortable answering.

For primary caregivers, questions about negative aspects of their child's behavior may be regarded by some as sensitive. It is important to view the child and teacher assessments in light of parental observations of children's behavior. Thus, parts of the Behavioral Assessment System for Children (BASC) are also included in the primary caregiver report. As part of the informed consent procedure, caregivers will be informed that they can refuse to answer any questions.

12. Estimates of Hour Burden

Exhibit 3 provides our estimate of time burden for the final data collection wave in Spring 2007 (school year 2006-2007) broken down by cohorts (Cohort 1 and Cohort 2). The table also indicates burden hours for the four waves of data that have already been completed (third grade fall, third grade spring, fourth grade fall, fourth grade spring). Exhibit 4 provides an overall

summary of the burden hours to respondents on each data collection instrument that will be administered during the 2006-2007 school year. Note that the multisite sample includes 100 schools rather than the initial 72 that was described in the original OMB submission. Appendix A2 and A3, provide

Exhibit 3 OMB Supporting Statement
 SACD Evaluation OMB Annual Burden Hours For School Year 2006-2007

Instrument	Cohort 1								Cohort 2								Totals
	Number of Respondents	Number of Responses per Respondent	Average Burden Hours per Respondent	Total Burden Hrs.	Data Collection School Year 2006-2007	School Year 2006-2007 Hours	School Year 2006-2007 Responses	School Year 2006-2007 Respondents	Number of Respondents ^a	Number of Responses per Respondent	Average Burden Hours/ Respondent	Total Burden Hrs.	Data Collection School Year 2006-2007	School Year 2006-2007 Hours	School Year 2006-2007 Responses	School Year 2006-2007 Respondents	
Child Report																	
Field Test	200	1	1.16	232	0	0	0	0	0	1	1.16	0		0	0		
Third Grade – fall	4243	1	0.83	3522	0	0	0	0	627	1	0.83	521		0	0		
Third Grade – spring	4447	1	0.83	3691	0	0	0	0	627	1	0.83	521		0	0		
Fourth Grade – fall	4600	1	0.83	3818	0	0	0	0	0	1	0.83	0		0	0		
Fourth Grade – spring	4600	1	0.83	3818	0	0	0	0	627	1	0.83	521	1	521	627	627	
Fifth Grade – spring	4600	1	0.83	3818	1	3818	4600	4600	0	1	0.83	0		0	0		
Teacher Report on Student																	
Field Test	18	11.11	0.58	116	0	0	0	0	0	0	0.58	0		0	0		
Third Grade – fall	335	12.67	0.25	1061	0	0	0	0	46	13.6	0.25	157		0	0		
Third Grade – spring	341	13.04	0.25	1112	0	0	0	0	46	13.6	0.25	157		0	0		
Fourth Grade – fall	340	13.6	0.25	1156	0	0	0	0	0	13.6	0.25	0		0	0		
Fourth Grade – spring	340	13.6	0.25	1156	0	0	0	0	46	13.6	0.27	169	1	169	626	46	
Fifth Grade – spring	340	13.6	0.27	1248	1	1248	4624	340	0	13.6	0.25	0		0	0		
Teacher Report on Classroom and School^f																	
Field Test	10	1	0.5	5		0	0	0	0	1	0.5	0		0	0		
Third Grade – fall	881	1	0.5	441	0	0	0	0	123	1	0.5	61		0	0		
Third Grade – spring	884	1	0.5	442	0	0	0	0	123	1	0.5	61		0	0		
Fourth Grade – fall	900	1	0.25	225	0	0	0	0	0	1	0.25	0		0	0		
Fourth Grade – spring	900	1	0.5	450	0	0	0	0	123	1	0.55	42	1	42	77	77	
Fifth Grade – spring	900	1	0.55	308	1	308	560	560	0	1	0.5	0		0	0		
Primary Caregiver Report																	
Field Test	200	1	0.58	116	0	0	0	0	0	1	0.58	0		0	0		
Third Grade – fall	3900	1	0.33	1287	0	0	0	0	545	1	0.33	180		0	0		
Third Grade – spring	3500	1	0.33	1155	0	0	0	0	545	1	0.33	180		0	0		
Fourth Grade – fall	4000	1	0.33	1320	0	0	0	0	0	1	0.33	0		0	0		
Fourth Grade – spring	4000	1	0.33	1320	0	0	0	0	545	1	0.33	180	1	180	545	545	
Fifth Grade – spring	4000	1	0.33	1320	1	1320	4000	4000	0	1	0.33	0		0	0		
SACD Activities – Principal Interview																	
Field Test	2	1	1.08	2	0	0	0	0	0	1	1.08	0		0	0		
Third Grade – fall	88	1	0.75	66	0	0	0	0	12	1	0.75	9		0	0		
Third Grade – spring	88	1	0.75	66	0	0	0	0	12	1	0.75	9		0	0		
Fourth Grade – fall	0	1	0.75	0	0	0	0	0	0	1	0.75	0		0	0		
Fourth Grade – spring	88	1	0.75	66	0	0	0	0	12	1	0.75	9	1	9	12	12	
Fifth Grade – spring	88	1	0.75	66	1	66	88	88	0	1	0.75	0		0	0		
SACD Activities – School Records																	
Third Grade – fall	88	1	4	352	0	0	0	0	12	1	4	48		0	0		
Third Grade – spring	88	1	4	352	0	0	0	0	12	1	4	48		0	0		
Fourth Grade – fall	0	1	4	0	0	0	0	0	0	1	4	0		0	0		
Fourth Grade – spring	88	1	4	352	0	0	0	0	12	1	4	48	1	48	12	12	
Fifth Grade – spring	88	1	4	352	1	352	88	88	0	1	4	0		0	0		
Total				34811		7112	13960	9676				2921		969	1899	1319	
Number of schools	88								12							100	
Annual Hours of Burden						7112	13960	9676						969	1899	8081	
Annual Responses							13960								1899	15859	
Annual Respondents								9676 ^a								1319 ^a	

^fEliminate double counting of teachers (Teachers who respond to the TRS also respond to the TRCS)

Exhibit 4
 Summary of Burden Hours Per Instrument for SACD Evaluation For School Year 2006-2007
 Burden in Hours to Respondents

Instrument	Number of Responses Cohort 1	Number of Responses Cohort 2	Total Number of Responses	Total Number of Respondents	Num of Responses/ Respondent	Average Burden Hours/Respondent	Total Burden Hours
Child Report	4600	627	5227	5227	1	0.83	4339
Teacher Report on Student	340	46	5250	386	13.6	0.27	1417
Teacher Report on Classroom and School	900	123	637 ^a	637 ^a	1	0.55	350
Primary Caregiver	4000	545	4545	4545	1	0.33	1500
SACD Activities – Principal Interview	88	12	100	100	1	0.75	75
SACD Activities – School	88	12	100	100	1	4.00	400
Total ^a	10016	1365	15859	10995			8081
Number of schools	88	12	100				

^aThis number has been adjusted downward to eliminate the double counting of teachers who completed both the TRS and the TRCS

documentation of the OMB approvals for this change in burden hours. The current change in burden hours (-7551) is described in section 15 below. The reduction is because there is only 1 data collection point for the current wave of data instead of 2 data collection points that occurred during 2005-2006.

As mentioned in the initial OMB submission, baseline data collection occurred among one cohort of third graders (cohort 1), associated teachers, school staff and caregivers beginning during the academic year 2004-05. During academic year 2005-2006, data collection pertaining to that same cohort of students while in the fourth grade took place. The final year of data collection on this cohort of students is scheduled to take place in Spring 2007 when this cohort will be in the fifth grade. During 2005-2006, a second cohort of students in 12 additional schools (4 sites) were surveyed to increase statistical power to detect meaningful impacts. Data was collected on this second cohort of students when they were in third grade (Fall 2005 and Spring 2006). In 2007, data collection will occur when these students are in the fourth grade (spring 2007).

The payments described above in Section A.9 compensate the respondents (parent/caregiver, and teacher) for the time they spend completing the data collection instruments, and thus there are no additional costs to respondents (parent/caregiver, and teacher) for the hours associated with the collection of information. School observations are not included in the burden estimate since the national evaluation staff will carry out this activity.

13. Estimate of Total Annual Cost Burden to Respondents or Record-Keepers

There are no direct costs to individual participants.¹⁴ Estimates of Annualized Cost to the Federal Government

The estimated cost to the federal government for the SACD Research Program National Evaluation—including designing and administering the baseline and follow up surveys, providing payments to respondents, processing and analyzing the data, and preparing reports summarizing the results—is \$7,634,028. The surveys and associated activities will be carried out over a four-year period. Thus, the average annual cost of the surveys and analyses is approximately \$1,908,507. This estimate is based on the evaluation contractor’s previous experience managing other research and data collection activities of this type.

15. Reasons for Program Changes or Adjustments

A program change of –7,551 has occurred. The reduction is because there is only 1 data collection point for the current wave of data instead of 2 data collection points that occurred during 2005-2006.

16. Plan for Tabulation and Publication and Schedule for Project

Our discussion of tabulation and publication plans focuses on the reports that will be produced after various rounds of follow-up data have been collected. We also discuss plans for tabulating descriptive information gathered from the baseline interviews and assessments that will be presented in these project reports.

a. Tabulation Plans

We will conduct three types of analyses to address the main impact-related research questions for the evaluation described in A1. First, we will conduct a *global* analysis to examine the extent to which the SACD initiatives improve elementary schoolchildren’s outcomes *overall*. This analysis will identify the particular social-emotional, school climate, behavioral, and academic-related outcomes that are most influenced by the SACD programs, and how overall impacts change over time. Second, we will conduct a *targeted* (or subgroup) analysis to

examine what works and for whom. In particular, we will examine whether impacts differ by key program features and structural elements, by baseline child and family characteristics, and by dosage level. Finally, we will conduct a mediated analysis to examine the pathways through which the interventions influence longer-term child outcomes.

Next, we discuss these analyses in more detail. The section begins, however, with a brief discussion of contextual analyses that we will conduct to aid in the interpretation of the impact estimates.

b. Contextual Analyses

The impact evaluation will begin with several contextual analyses that will lay the foundation for the impact analysis, and that will be crucial for interpreting the impact results.

These analyses include:

1. Assessing how well random assignment was implemented to examine the extent to which the impact estimates (treatment and control group differences) are unbiased
2. Assessing how the new entrants to the study in the follow-up waves of data collection (i.e. new students who enter the research schools) can be included in the impact analyses
3. Examining the baseline characteristics of children in the treatment and control schools to understand the student population under investigation
4. Examining the social and character development services received by treatment and control group members to understand the nature of the SACD interventions offered in the treatment schools and the counterfactual for the evaluation

Assessing the Integrity of the Random Assignment Process. The generalizability, validity, and interpretation of the impact estimates hinge on the integrity of the random assignment process and adherence to its procedures. We will conduct several analyses to gauge the success of the random assignment process. First, we will examine the mobility of children in the sample into and out of the treatment and control schools using follow-up interview and

school records data. Such movers complicate the analysis, because to preserve the integrity of the random assignment design, children who relocate from treatment to control group schools must be considered treatment group members in the analysis, and similarly, children who relocate in the reverse direction must be considered controls. We will use statistical procedures to account for these crossovers.

Second, within and between sites, we will conduct statistical tests, using baseline data and school records data covering the pre-intervention period, to gauge the similarity of the baseline characteristics of students in the treatment and control schools. We expect that the random assignment and pairwise matching processes used to select the schools in the research sample will produce equivalent treatment and control groups.

Finally, we will monitor changes, *unrelated* to the SACD interventions, that could affect student- and school-level outcomes and student mobility in the communities where the treatment and control group schools are located. These events might include unexpected changes in employment prospects (such as the closing of a large plant), changes in the crime rate, changes in school policies (such as the introduction of a zero tolerance policy), or changes in school or district personnel (such as a new principal or superintendent). These changes could lead to biased impact estimates if they are not controlled for in the analysis. Information on these events will be collected through discussions with the grantees.

Assessing How the New Entrants Can Be Included in the Impact Analyses. Once new entrants are included in the evaluation, we must consider how this sample can be used in the multisite impact analysis.¹ This is a complex issue because there may be differences in the observed and unobserved characteristics of new students who enter the treatment and control

¹For purposes of estimating burden (Exhibit 3), we have assumed that new entrants will be included in all data collection activities throughout the course of the study. Thus, we have provided a “maximum” estimate of burden.

group schools during the post-random assignment period. These differences could result from the interventions themselves, if, for example, families decide to move into the areas served by the treatment group schools because they want their children to be exposed to the SACD interventions. Alternatively, differences between the treatment and control group refresher samples could result from factors unrelated to the interventions (such as changes in local-area employment prospects, changes in the local crime rate, or turnover of school staff). These factors could significantly alter the composition of students in the treatment or control group refresher samples due to the relatively small number of schools per district that are included in the study.

If the average characteristics of new entrants differ across the treatment and control groups in ways that are correlated with key student outcomes, it would be difficult to interpret impact estimates that are based on samples that include the new entrants. This is because the impact estimates would confound two effects:

1. The extent to which the SACD interventions improve the outcomes of the *average* student in the district (at the time of random assignment)
2. Differences in the average outcomes of treatment and control group students due to differences in the composition of students that enter the two types of schools during the post-random assignment period (due to factors either related or unrelated to the interventions)

For example, if students with high test scores are more likely to move into the treatment than control group schools, it would be difficult to determine whether positive impacts on test scores were due to the SACD interventions or to the possibility that treatment group members would have had higher test scores even if they had not been exposed to the interventions.

If the new entrants appear to differ systematically across the treatment and control groups, the main impact analysis will be conducted using only the original sample members. We will, however, conduct supplementary analyses that include the new entrants in order to examine the

robustness of study findings, but we will interpret these results carefully. We will also carefully document the characteristics of the new entrants in both the treatment and control group schools to help interpret the main impact estimates, because the outcomes of original sample members may be influenced by the new entrants.

If the new entrants in the treatment and control groups appear to be comparable, based on their observable characteristics at the time they enter the study schools, we will consider pooling the refresher students with the original sample for the impact analyses. The inclusion of these new entrants will increase the precision of the impact estimates relative to those based on original sample members only.²

Students in the refresher sample will have been *exposed* to the intervention for less time than students in the original sample. Thus, including these students in the impact analysis may dilute estimated program impacts. For this reason, we also will estimate impacts using the original sample of students only, which will allow us to assess the impact of the program on the set of students with the same potential exposure to the program. As discussed later, however, the refresher sample will play an important role in the analysis of dosage effects.

Examining Sample Characteristics. We will conduct comprehensive descriptive analyses of the characteristics of the sample to help us more fully understand the types of children and families in the research sample, including their backgrounds and risk factors. These results will help us interpret program impact estimates, and guide us in defining subgroups that may be of policy interest. These analyses will be conducted using baseline interview and assessment data as well as school records data. In addition, geographic information will be linked to sample

²We will carefully note that, although the observable characteristics of the treatment and control group refresher samples appear to be similar, their unobservable characteristics may differ, which might yield biased impact estimates.

members (by zip code or county) to examine the characteristics of the communities in which sample members live.

As part of this descriptive analysis, we will use national data (for example, data from the Early Childhood Longitudinal Survey [ECLS]) to examine how our sample of third graders compares to nationally and locally representative samples of third graders. These analyses will help us assess the generalizability of our findings.

Finally, if response rates to the consent form are lower than expected, we will need to assess whether the children who consented to participate in the study are representative of all third-graders. This analysis will help us assess the generalizability of the impact findings. Specifically, we will use school records data to compare average grades, absences, tardies, and test scores of the consenting children to those of all third graders. If the groups are similar on these variables, impact estimates based on the sample of consenting children are likely to be generalizable to the full set of third graders. If the two groups have different characteristics, however, then the impact estimates may not be generalizable to the full set of third graders.

Examining the Receipt of Services Targeting Social and Character Development. To understand estimates of the impact of the SACD interventions on child behavioral and academic-related outcomes, it is crucial to understand the intensity and nature of the social and character development services received by *both* treatment and control group children in each site. We can expect beneficial impacts of the SACD interventions only if treatment group children receive well-implemented and well-designed SACD program services, and the size of the impact is likely to be correlated with the amount, intensity, and quality of services received. Similarly, it is crucial to obtain information on the social and character development services offered to control group children, because the evaluation is assessing the effectiveness of the SACD programs relative to the status quo curriculum in the school districts, which might include *other*

SACD-like programs. Thus, information on services received by control group children is needed to define the “counterfactual” for the evaluation. This descriptive analysis will be conducted using data from the SACD-Activities Observation Instrument.

c. Construction of Student And School Outcomes

We will use interview and assessment data to construct outcome measures in four domains: (1) social-emotional competence, (2) school climate, (3) students’ behavior, and (4) students’ academic achievement. Many of these outcome measures will be based on scale scores, whereas others (for example, some academic achievement measures) will not be scale-based. Next, we briefly outline our approach for constructing these two types of outcome measures *consistently* across all sites.

Outcomes Based on Scales. Our goal is to create scale-based outcome variables for the impact analyses that reliably measure distinct constructs. Our general approach to doing this will involve the following steps:

- ***Constructing outcome measures from scale items according to scale developers’ instructions.*** We will consult published materials and contact test developers to obtain detailed instructions for constructing scales. We will modify the test developer’s instructions if a trend has emerged in the literature for analyzing the scale differently. We will follow the developers’ practices for handling missing item-level data or if necessary, establish criteria for how much missing data is acceptable and whether missing items will be imputed.
- ***Assessing the distribution and reliability of the outcome measures.*** We will examine the distributions of the constructed outcome variables and calculate Cronbach’s alpha for each scale and its subscales. If an alpha is low, we will examine how well each item coheres with the others and conduct factor analyses to identify whether items should be dropped or (different) subscales should be constructed to create outcome measures with adequate reliability.
- ***Assessing whether these outcome measures tap distinct constructs.*** We will conduct exploratory analyses of the data to detect underlying latent factors that may better serve as outcome variables in the impact analyses. If the outcome measures created from previously defined scales do not tap distinct constructs, these analyses may yield new constructs. We will assess the distribution and reliability of the new constructs. We will use a random half of the data to conduct exploratory analyses

using factor analyses and structural equation models. Once we have developed constructs with the best properties, we will conduct confirmatory analyses using the other random half of the data.

Outcomes Not Based on Scales. In addition to the scale construction described above, we will construct variables for other "non-scale" outcomes, which measure students' academic achievement and behavior. For example, we will be gathering information on students' grades, attendance, test scores, and school disciplinary actions, and we will use this information to construct variables consistently across all sites for use in the impact analysis. For grades, we plan to convert letter grades to numeric grades based on grading scales collected from schools, presenting impacts on grades in each subject. For test scores, we plan to present impacts on test scores in percentile units. For attendance, we plan to create variables for number of days absent and percentage of students absent, and for behavioral outcomes we plan to construct variables for the percentage of students suspended and the number of times suspended. We will also create binary variables signifying whether a student has particularly poor outcomes (for example, whether the student has test scores or attendance levels below pre-specified cutoff values). We will also investigate the feasibility of constructing other measures depending on what measures are available from school records.

d. Global Analysis

The global analysis will examine the extent to which the SACD interventions, *on average*, change children's outcomes relative to what these outcomes would have been otherwise. Although, as discussed, the SACD interventions differ across sites, it is of policy importance to examine the overall effectiveness of the SACD initiatives, to examine which particular outcomes have the largest overall impacts across all sites, and how impacts change over time.

Basic Statistical Model to Estimate Point-in-Time Impacts. Random assignment of schools will be performed before children enter third-grade. Thus, unbiased estimates of the impacts of the *offer* of the SACD interventions (relative to other program alternatives offered in the control schools) can be computed as the differences in the average outcomes of all treatment and control group children. This approach yields unbiased estimates of the “intention-to-treat” impacts, because the random assignment design ensures that the main difference between the treatment and control groups at the point of random assignment is the opportunity to receive SACD program services.³

Although we will compute these simple differences-in-means impact estimates, we will focus on regression-adjusted estimates. This is because regression procedures improve the precision of the estimates, and adjust for residual differences in the observable characteristics of program and control group members due to small sample sizes, random sampling, and interview nonresponse.

We will estimate regression-adjusted impacts using hierarchical linear methods (HLM), because this approach accounts for the nesting of children within classrooms and schools. The basic model consists of three levels that are indexed by children (*i*), classrooms (*c*), and schools (*s*). The three levels can be aggregated into a unified model, which in its simplest form, can be expressed as follows:

$$(1) \quad Y = \sum_j \beta_j(\text{Site}_j * T) + X\gamma + Y_0\delta + Z\lambda + [\theta_s + \eta_{cs} + e_{ics}],$$

where

³The random assignment of only 5 schools per site to the treatment group and 5 schools per site to the control group may lead to some differences between the demographic and background characteristics of children in these two types of schools. However, as discussed later, the pairwise matching process that will be used to select the research sample will minimize these differences to the greatest extent possible.

- Y = Child outcome at a specific follow-up time point, such as the self-efficacy scale or standardized test scores
- $Site_j$ = Indicator variable equal to 1 if the child is in site j , and 0 otherwise
- T = Treatment indicator equal to 1 if the child is assigned to the treatment group, and 0 if the child is assigned to the control group
- X = Child and family demographic characteristics pertaining to the period prior to random assignment, such as child's gender, race/ethnicity, and family income
- Y_0 = Baseline measures of the outcome measures (from the fall 2004 interviews and assessments), such as child test scores, child aggression scores, and primary caregiver prosocial behavior scores
- Z = Baseline aggregate school measures (or indicators of school pairs) used in the matching process⁴
- $\beta, \gamma, \delta, \lambda$ = coefficients to be estimated
- $\theta_s, \eta_{cs}, e_{ic}$ = random (and mean zero) school-level, classroom-level, and individual-level error components (effects), respectively

In words, equation (1) says that any given child outcome at a point in time is determined by the child's baseline level of development, his or her family background, aggregate school characteristics, the intervention (in this case the opportunity to receive SACD services), and a set of other factors that are not related to his or her intervention assignment status. In this formulation, the estimate of β_j represents the regression-adjusted impact estimate for site j .

We highlight several important features of the regression model. First, because random assignment will occur at the school level and not at the student level, the model incorporates the *clustering* of students within schools and classrooms (which reduces the precision of the estimates). Second, in the analysis we will give each site (grantee) equal weight regardless of sample sizes within the sites. The SACD interventions will be administered at the site level and will differ across sites; thus, the site is the relevant unit of analysis. Accordingly, impact

⁴Because of the relatively small number of schools and classrooms in the sample, only a small number of these measures can be included in the models to avoid model overfitting.

estimates across all sites will be obtained by taking the simple average of the regression-adjusted impacts in each site (that is, the β_j s). The associated t-tests will be used to test the statistical significance of the impact estimates.

The explanatory variables included in the regression models will be obtained from the Fall 2004 measures. We expect that the explanatory variables will substantially increase the precision of the impact estimates, because some (and in particular, the baseline measures) are likely to be highly correlated with the outcomes measured at follow-up.

The statistical methods used to estimate the regression models will depend on the nature of the outcome measure. For example, we will use ordinary least squares methods for continuous outcome measures (such as test scores, attendance, or aggression scale scores), and logit maximum likelihood methods for binary ones (such as the percentage of children with low test scores or low school attendance).

Finally, equation (1) can be used to estimate impacts on outcomes measured at the entire *school* level (for example, school climate measures). For these analyses, the dependent variable, Y , will be measured at the school rather than at the child level and the explanatory variables will include only baseline school characteristics. Furthermore, the error structure will include only random school effects. Similarly, for analyses examining intervention effects on *teacher* outcomes (for example, the teacher involvement scale), the dependent variable will be measured at the teacher (classroom) level.

Longitudinal and Growth Curve Models. A major strength of the SACD evaluation is the measurement of child outcomes at *five* time points: Fall 2004, Spring 2005, Fall 2005, Spring 2006, and Spring 2007. This presents an opportunity to learn about both short- and medium-term impacts. We will estimate impacts over time using various approaches. First, the regression model in equation (1) will be estimated for each time period separately.

This analysis will generate period-by-period impact estimates. Second, we will extend the model in equation (1) to estimate period-specific impacts *simultaneously* to obtain more precise impact estimates. Specifically, we will estimate longitudinal models in which outcomes across time periods are stacked. Third, we will estimate program impacts using *growth curve model techniques*. These models will be used to examine impacts (treatment and control group differences) on the growth trajectories of child outcomes during the follow-up period. We will examine and compare results obtained using the various interrelated statistical approaches.

The longitudinal model allows us to examine efficiently three key hypotheses for outcomes in each domain:

1. **Steady Impacts:** Gains made in Year 1 continue through the end of grade 5
2. **Fadeout:** Gains made in Year 1 shrink or disappear by the end of grade 5
3. **Delayed impacts:** Gains start to show up in Years 2 or 3

e. Targeted Analysis

The targeted analysis will use a more refined approach than the global analysis to examine the effects of the SACD interventions on key child-, teacher-, and school-level outcomes. The targeted analysis will address the important policy questions of what works, and for whom. Specifically, this analysis will address the extent to which impacts vary across key program characteristics and according to key child and family subgroups. The analysis will also examine if impacts differ by the amount and intensity of intervention that is received. The results of these analyses have important policy implications, both for the operation of the SACD programs and for the future program development of other similar initiatives.

Subgroups Defined by Program Characteristics. Impact results by key structural elements and features of the SACD interventions can provide important information on how to

improve program services, as well as to develop and expand the programs targeting social and character development among elementary schoolchildren.

The program-related subgroups will be determined in consultation with IES, CDC, and the grantees after the SACD interventions have been implemented, and after descriptive data have been collected on the nature of the interventions (from the SACD activities instrument). Subgroups will be identified that are policy relevant and that reflect important dimensions of program variation. Because of relatively small evaluation sample sizes, we will estimate impacts for only a small number of key subgroups for whom relatively precise impact estimates can be obtained.

We expect that the final list of subgroups will include those in the following categories:

- **Program structure**, including whether the primary targeting unit is the classroom, entire school, or another entity (an afterschool program or the family), and whether the SACD curriculum consists of distinct activities or is embedded in the regular curriculum.
- **Curriculum content**, including, for example, whether the primary SACD curriculum focuses on social skills training, behavior modification, or values clarification.
- **Dosage and intensity of the intervention**, including the number of hours and days per week the intervention is offered.
- **Quality of program implementation**, including a categorical scale depicting the level of fidelity to the program model and the quality of services provided.⁵

The random assignment design allows us to estimate unbiased estimates for sites with a specific program characteristics by comparing the outcomes of treatment and control group members in those sites. For example, we can obtain unbiased estimates for sites with a high quality service environment by estimating the regression models using treatment and control group members in those sites. The models can also be used to test, for example, whether

⁵Grantees are developing measures of fidelity and will work together to determine how best to analyze them.

impacts are larger in sites with well-implemented programs than in other sites, or whether impacts are larger in sites whose programs target the entire school rather than the classroom.

We will also use hierarchical linear models (HLM) to help isolate the effects of particular program features from others. The HLM models will be estimated in two stages. In the first stage, we will obtain impact estimates for each *site* using equation (1). In the second stage, we will estimate the following model, where the site-specific impacts are regressed on key measures pertaining to the program subgroups (denoted by W):

$$(3) \quad \hat{\beta}_i = \alpha + W\gamma + u,$$

where α and γ are parameters to be estimated and u is a mean zero error term. The results from these models can be used to disentangle the effects of particular program features from others. Because of the relatively small number of sites included in the evaluation, however, we will only be able to include a few key site characteristics in the regression models in order to avoid “overfitting” the models.

Subgroups Defined by Child and Family Characteristics. An important policy issue is the extent to which the effects of alternative social and character development initiatives vary across children with different background characteristics. We will use baseline interview and school records data to define key child subgroups across which program effects might vary. Although the final list of subgroups will be selected in consultation with IES, CDC, and the grantees, we expect them to include:

- ***Child and family demographic characteristics***, such as child’s gender, race and ethnicity, family’s poverty status, and risk status (constructed using cluster analytic techniques to obtain a single measure summarizing key risk factors faced by the child)
- ***School history measures***, including prior test scores, grades, and attendance.

- **Child behavior measures**, including key baseline social-emotional competence and behavior measures
- **Parenting measures**, including key family moderators such as parenting practices and the home atmosphere

We will obtain these subgroup impact estimates using procedures very similar to those described above for the program-related subgroups. We will estimate equation (1) to compute regression-adjusted impacts for children in a particular subgroup. For example, we will estimate impacts for boys by comparing the mean outcomes of boys in the treatment and control groups. In addition, we will conduct statistical tests to gauge the statistical significance of the subgroup impact estimates, and the differences in impacts across levels of a subgroup (for example, for boys and girls). We will include also child subgroup indicator variables in the HLM models to help disentangle child subgroup effects from program ones.

Estimating Dosage Effects. For several reasons, we expect differences in the amount and intensity of SACD intervention services that are received by treatment group children. First, dosage levels will differ across the SACD program models. Second, some children in the research sample will leave the treatment group schools, and, hence, will receive fewer SACD services than those who remain longer in these schools. Third, children who are added to the sample after fall of 2004 will be exposed to the program for less time than children from the original sample. Finally, school attendance will differ across children, which could lead to differences in exposure to the SACD interventions. Thus, an important research question for the evaluation is: Are impacts larger for children who receive a higher dose of the treatment than for those who receive a lower dose?

We will use a variety of statistical procedures to estimate program impacts for SACD participants who receive varying amounts of SACD services. First, we will compare impact estimates in sites that offer intensive SACD program services with those in sites whose program

curricula are less intensive. These findings, which are fully based on the random assignment design, will provide some evidence of the extent to which impacts differ by dosage level (although there could be other site-specific factors that could contribute to differences in impacts across sites).

Second, we will carefully examine changes in estimated impacts over time. Evidence of increases in impacts over time will be suggestive of the presence of dosage effects (since dosage levels will increase over time), although changes in impacts over time could also result from other factors (such as delayed program effects). Again, this approach has the advantage that it relies on the random assignment design. We will test also the robustness of these findings by estimating impacts over time using only treatment and control group members who remain in their schools for the entire period, if the baseline characteristics of these two groups of children appear to be similar.

Third, as new entrants are added to the sample during the follow-up period (and if the characteristics of these new entrants are similar in the treatment and control group schools), we will examine dosage effects by comparing impact estimates for new entrants with those for original sample members. Because the new entrants will have been exposed to the interventions for a shorter amount of time than the original sample members, we might expect that impacts for the new entrants will be smaller. However, these results will need to be interpreted carefully, because other factors could influence the relative sizes of the impact estimates for the refreshed and original samples. For instance, if the quality of program implementation improves over time, then, for a given level of program exposure, the new entrants might benefit more from the interventions than the original sample (that is, impacts after one year of exposure might be greater for the refreshed sample than for the original sample). As another example, the characteristics of new entrants and original sample members might differ, which could influence

the size of the impact estimates if program effects differ across child subgroups. Nonetheless, the use of the refresher sample to help tease out dosage effects will be an important component of the dosage analysis.

A fourth, and more general approach, that we will use to estimate dosage effects will be to use *propensity scoring* to match treatment group members in a particular dosage group to control group members with similar baseline characteristics. Dosage effects will then be estimated by comparing the outcomes of treatments in a particular dosage category to their matched controls. For instance, we will estimate the effects of SACD services for those in the high-dosage group, by comparing the distribution of outcomes of high-dosage treatments to their matched controls. Similarly, we will estimate the effects of SACD on those who receive less of the intervention by comparing the outcomes of low-dosage treatments to their matched controls.

Finally, in order to test the robustness of our findings using the propensity scoring approach, we will also estimate dosage effects by (1) calculating, for each *treatment* group member, the difference between their outcomes in the follow-up period relative to their corresponding baseline outcomes (that is, the growth in their outcomes), and (2) comparing the mean difference in these growth rates for those that received different amounts of the intervention. This “fixed-effects” or “difference-in-difference” approach adjusts for selection bias by assuming that permanent unobservable differences between children across dosage groups are captured by their baseline (pre-intervention) measures.

f. Mediated Analysis

The SACD interventions aim to influence children’s behavior and academic achievement both directly and indirectly through their effects on the school climate and children’s social-emotional competence. The analyses described so far, however, have not addressed the

mechanisms whereby some mediating outcomes ultimately influence more distal child outcome measures.

Thus, we will also conduct a mediated analysis to examine these mechanisms. The analysis results can be used to examine whether impact estimates for the evaluation are internally consistent (that is, “make sense”) based on the theoretical relationships between mediating and longer-term outcomes. Second, program staff can use the analysis results to focus efforts on improving mediating behaviors that SACD interventions have large impacts on and that are highly correlated with longer-term child outcomes.

The approach to the mediated analysis can be considered a three-stage process. In the first stage, a longer-term outcome measure is regressed on mediators and other explanatory variables (moderators). In the second stage, the regression coefficient on each mediator is multiplied by the impact on that mediator. These products—labeled “implied impacts”—are what we would expect the impacts on the longer-term outcome to be on the basis of the relationship between the mediators and the longer-term outcome. Finally, the *implied* impacts are compared to the *actual* impacts on the longer-term outcome. These results indicate the extent to which impacts on the longer-term outcome variable can be partitioned into impacts due to each mediator.

We will use the conceptual model discussed above to specify the models that will be tested. For example, we will examine the associations between impacts on children’s social-emotional competence measures (such as social problem solving, attitudes about aggression, self-efficacy, and empathy) and the impacts on behavioral outcomes (such as altruistic behavior, aggression, minor delinquency, disruptive classroom behavior, and victimization). The choice of specific models to be tested, however, will not be based solely on theoretical considerations, but also on the empirical findings. For example, it will only be meaningful to conduct mediated analyses using mediators or longer-term outcomes that are shown to be significantly influenced by the

SACD interventions. Furthermore, many of the mediators may be highly correlated with each other, making it difficult to isolate the effects of some mediators from others. In this case, we will carefully select appropriate measures to include in the mediated analyses to obtain meaningful results.

Publication Plans. Three major evaluation reports will be published highlighting findings from each year of SACD program implementation (2003-2004, 2005-2006, 2006-2007). The reports are scheduled to be completed in March 2007, January 2008, and September 2008, respectively (see Table 1). A key objective of the reports will be to discuss the impacts of the program on student outcomes. Outcomes at the school-level, such as measures of school climate, will also be examined to assess whether the programs are having positive effects on the overall atmosphere of schools participating in the program. Findings from the contextual analyses discussed above will also be part of these reports.

Time Schedule. The full timeline for the evaluation is shown in Table 1. Major design and school selection activities occurred between October 2003 and September 2004. Data collection for the baseline year occurred between September and December 2004, with follow-up interviews and assessments occurring in spring 2005, fall 2005, and spring 2006. The fourth follow-up data collection is scheduled for March 2007-June 2007. To date, baseline, first, second, and third follow-up data collection activities have been completed.

TABLE 1
SCHEDULE OF ACTIVITIES

<u>Activity</u>	<u>Schedule</u>
Design and sample selection	October 2003-December 2004

Baseline data collection	September 2004-December 2004
First Followup	March 2005-June 2005
Second Followup	September 2005-December 2005
Third Followup	March 2006-June 2006
Fourth Followup	March 2007-June 2007
Reports	
First Impact Report (Third Grade Year)	March 2007
Second Impact Report (Fourth Grade Year)	January 2008
Third Impact Report (Fifth Grade Year)	September 2008
Technical Memos and Brief Reports	As requested

17. Approval Not to Display the Expiration Date for OMB Approval

Approval not to display the expiration date for OMB approval is not requested.

18. Exception to the Certification Statement

No exceptions to the certification statement are requested or required.