ABSTRACT

Title of Dissertation:	COST-EFFECTIVENESS OF AN ENHANCED WHOLE-SCHOOL SOCIAL COMPETENCY INTERVENTION
	Sharon Huang, Doctor of Philosophy, 2008
Directed By:	Professor Gary D. Gottfredson Counseling and Personnel Services

Background

The cost and efficacy literature regarding elementary school-based preventative programs is limited, and many cost and efficacy studies suffer from research design and methodology problems.

Purpose

This study compares the marginal costs of a specific whole-school intervention to marginal costs of control school programs. It also assesses the cost-effectiveness of treatment and control interventions with respect to self-reported aggression, academic grades, and Maryland State Assessment (MSA) scores.

Setting

The study takes place in elementary schools in Anne Arundel County, a relatively diverse, suburban county in Maryland.

Subjects

Counselors ($n=9$) and principals ($n=11$) from 12 schools; the project manager, data clerk, and coordinator of guidance from the school system central office; and five research team members completed questionnaires about their time use. The study relies on efficacy data from another study.

Intervention

Implemented for three years, Second Step is a popular preventative, school-wide social competency program that aims to augment students' social skills and prevent problem behavior (Frey, Hirschstein, & Guzzo, 2000). Using manual-based lessons, classroom teachers in first through fifth grades were trained to deliver 30-minute lessons once a week to their classes in the areas of empathy, anger management, problem solving, and impulse control. Teachers also supplemented the formal lessons by reinforcing what had been taught at other times during the day (Frey et al., 2000). The intervention was enhanced by adding specific implementation standards and the use of periodic feedback about implementation to intervention managers and teachers.

Research Design and Methods

This study makes use of results from a large-scale randomized controlled trial that investigated the efficacy of Second Step. Researchers selected 12 elementary schools that had never implemented Second Step to participate, and these schools were matched based on their demographics and achievement history. Within each pair of schools, researchers randomly assigned one school to the treatment group and one to the control group. The methodology used for assessing costs is the ingredients approach (Levin & McEwan, 2001).

Data Collection and Analysis

Effect sizes for third and fourth graders for the third year are obtained from outcome evaluation reports. Third-year costs are added to retrospective training costs to estimate three-year costs. Data for costs of personnel time are collected in the form of time-use questionnaires, supplemental teacher questionnaires, and implementation logs.

Accounting expenditures, rental agreements, and contacts with district personnel provide other cost data.

To reflect current and annual costs and to account for opportunity costs, costs are discounted (expressing future costs in terms of their present values) and amortized (distributing a cost across its lifetime). Finally, incremental cost-effectiveness ratios are calculated for some of the outcome measures examined. Sensitivity analyses are conducted to consider variability in cost and cost-effectiveness estimates.

Results

Results imply that the enhanced whole-school social competency intervention attained no positive effects in student self-reported aggression, academic grades and MSA scores, at a cost of only $69 less per student over a three-year implementation period.

Conclusions

The enhanced whole-school social competency intervention is no more cost-effective than the control programs: annual marginal student costs are only slightly less than those of the control program, and the program is not efficacious. Limitations include the small number of schools and personnel and correspondingly large standard errors for effect sizes, the use of self-report methods to estimate time, and dependence on unreliable accounting expenditure data from the school district. Undependability of cost data may result from both measurement error and bias. In addition, since this particular Second Step program was implemented in only one school district, its generalizability to other school districts or variations of program implementation is unexplored.

Despite these limitations, the study provides a range of credible values for cost-effectiveness for the program. It may provide insight to the scientific community about the costs involved in operating an enhanced whole-school intervention to share with school administrators and educators in their considerations of elementary school-based preventative interventions.

References

Frey, K. S., Hirschstein, M. K., & Guzzo, B. A. (2000). *Second Step*: Preventing aggression by promoting social competence. *Journal of Emotional and Behavioral Disorders, 8*, 102-112.

Frey, K. S., Nolen, S.B., Edstrom, L.V.S., & Hirschstein, M.K. (2005). Effects of a school- based social-emotional competence program: Linking children's goals, attributions, and behavior. *Journal of Applied Developmental Psychology, 26*, 171-200.

Levin, H.M. & McEwan, P.J. (2001). *Cost effectiveness analysis: Methods and applications*. Thousand Oaks, CA: Sage.

Cost-Effectiveness of an Enhanced Whole-School Social Competency Intervention

By

Sharon Huang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2008

Advisory Committee:

Professor Gary Gottfredson, Chair
Assistant Professor Randi Hjalmarsson
Professor Steven Klees
Professor Sylvia Rosenfield
Associate Professor William Strein

Acknowledgements

Table of Contents

List of Tables

Cost-Effectiveness of an Enhanced Whole-School Social Competency Intervention

School systems should adopt effective preventative and early intervention programs (Aos, Lieb, Mayfield, Miller & Pennucci, 2004). These programs should be implemented well. But, while schools may select potentially beneficial programs and implement them well—or implement well interventions that bring no potential benefits— the decision on which school programs to select and implement often rests on costs of these programs. Even if schools recognize the costs of their programs, they may not have information on whether the long- and short-term benefits to students served are worth these costs. Similarly, even if information about the benefits of programs is available, information about benefits relative to costs may not be. The scientific community needs information on cost-effectiveness of elementary school-based preventative programs to advise administrators and educators about potential programs to select for their schools.

Cost-effectiveness studies in education are usually conducted to provide information about current and future costs and outcomes of alternative programs, to assist in decision-making about the future. In addition, these studies offer insight about how resources should be invested to attain a desired outcome. With this knowledge, the scientific community may be able to inform administrators and educators about selecting, starting, continuing, and institutionalizing programs that will maximize school resources while enhancing school outcomes.

The purpose of this study is to compare the cost-effectiveness of an enhanced, school-wide, social competency intervention to that of control school social competency interventions within a local school district. Hence, the study adopts the perspective of the school district, which is likely to be interested in costs and outcomes of the evaluated

intervention. This information serves to assist the school district in addressing policy

concerns about whether or not to expand the intervention to additional schools. Costs of

treatment and control interventions are assessed with respect to three outcomes reported

by an outside outcome evaluation: self-reported aggression, academic grades, and

Maryland State Assessment (MSA) scores. Cost-effectiveness is expressed as the average

marginal cost of the intervention per student per standard deviation unit gained or

reduced in the outcome. The methodology used for measuring costs is the ingredients

approach (Levin & McEwan, 2001), which is explained in Table 1 and in the

methodology section. To assist the reader of this paper, Table 1 defines several other

common ideas and approaches encountered in cost analysis studies.

Table 1

Terms Used in Studies of Program Cost

Term	Definition and Explanation
opportunity cost	The value of an alternative that is given up when it is used in another way. For example, if fifth grade students miss half an hour of social studies to tutor first grade students, the opportunity cost is the loss of the fifth grade students' instructional time in social studies.
ingredients approach	A common method for estimating the total costs of a program or intervention. The evaluator must thoroughly identify all the resources involved in implementing the program or intervention and place values on them.
market prices	The values of goods or services offered in a competitive market. Market prices are frequently used to estimate costs of resources.
discounting	Requires the assumption that costs that arise in the future would be less than costs incurred in the present, in other words, a dollar in the present is worth more than a dollar in the future, economists use discounting to adjust future costs to present costs. Future costs must

	be discounted to make comparisons between future and present costs. Thus, discounting requires the assumption of a *discount rate* that expresses the future costs of something in terms of its present value. The higher the discount rate, the lower the value of the future costs is in present terms.
amortization or annualization	The distribution of the costs of an item over its lifetime. Because a resource, such as equipment, is typically used beyond one year, its cost must be prorated over the years of its lifetime.
sensitivity analysis	A technique for determining a range of cost estimates, to account for uncertainty in the cost analysis. Considering a range of values for a parameter—for instance, discount rate—may reveal differences in overall costs.

Conceptual Framework for Costing

The issue of cost-effectiveness in starting an enhanced social competency intervention in additional schools within a school district provides the framework for measuring costs in the present study. The social competency intervention would be implemented for a period of three years, as recommended in the cost literature. To answer this policy question, the annual costs of the social competency intervention are assessed over its initial three years of implementation. Due to availability of data, marginal program costs are collected in the third year. To approximate first- and second-year costs, annual program costs calculated from the third-year data are added to the estimated costs for startup materials and personnel training, incurred largely in the first two years of program implementation. To project the total three-year costs of the social competency intervention in new schools, future costs are converted back to first-year (present) costs using a range of discount rates recommended by cost literature. This conversion using discount rates allows for comparison of costs across different points in time.

Apart from adjusting for future costs using the concept of discounting, accounting for the lifetime of resources by using annualization when necessary is also integral to a cost-effectiveness analysis. Resources can be regarded as having either capital or recurrent costs. Capital costs describe costs of resources that are generally useful for more than one year, and are therefore annualized. The process of annualization allows for just the portion of the resource and amount of time utilized during the intervention to be counted in total annual intervention costs. By contrast, recurrent costs describe costs of resources consumed within a year, and therefore do not require annualization. In the current cost-effectiveness analysis, initial trainings, facilities, materials, and equipment are considered capital costs, while refresher trainings are considered recurrent costs.

Two key processes are used to account for comparison of costs incurred across different points in time: discounting and amortization. The following sections present an overview of these two concepts, which account for opportunity costs and are widely used in cost analyses.

Discounting. Discounting considers society's preference for acquiring benefits in the present and incurring costs in the future, suggesting that future costs are worth less than present costs. As a result, the further away in the future costs are incurred, the more these costs are discounted, as indicated by a denominator increase in the discount factor formula below. The discount rate is an interest rate used to convert the future costs of a resource into its present value, as a means of comparing costs that occur at different points in time. The equation below calculates the present value of future costs (Levin, 1983), which takes into account the time at which costs are incurred (t) and the discount rate (r):

$$PV = \sum_{t=1}^{t} \frac{C}{(1+r)^{t-1}} \qquad (1)$$

The net present value of costs in year *t* is the sum of annual discounted values. Discounting considers the opportunity costs of the undepreciated part of the resource, which may have been invested in an alternative opportunity and earned interest. However, the present value of the resource varies from year to year, even though the resource is most likely consumed at the same rate each year; hence, amortization may be used instead of discounting to reflect the equivalent annual costs of the resource (Walker & Kumaranayake, 2002).

The choice of discount rates for this study is based on recommendations of cost analysis literature, consideration about the projected return on the investment of the enhanced whole-school intervention compared to its alternative, and current economic factors such as inflation and interest rates. For educational programs, the literature recommends using low discount rates. One approach to determining the discount rate is consideration about the average expected returns on investment from investing financial resources in an alternative opportunity (Levin & McEwan, 2001). For example, if the financial resources required for Second Step implementation could return five percent if invested in an alternate program, a five percent discount rate should be selected. Given current economic factors, there would be a lower expected return on investment. Thus, the selection of discount rates reflects opportunity costs for the school district that will benefit from the returns on its investment. Sensitivity analyses are included to show differences in net present value due to variations in discount rates. Considering all of the above factors, this study uses discount rates of three and five percent. Figure 1 provides an example of how costs are discounted by time.

5

Figure 1. Discount rate example

Two different school programs are implemented over a period of five years. Program 1 requires an up front cost of $5,000 in the first year, while program 2 requires an annual cost of $1,000. To compare the costs of the two programs, discount rates are used to determine the present value of program 2, since its costs occur in the future. At a discount rate of 3%, the present value of the cost for program 2 would be $4,717. However, at a discount rate of 5%, the present value of the cost for program 2 would be $4,546. A higher discount rate may be used to account for the fact that program 2 may reflect a riskier investment and may depreciate more quickly. Compared to program 1's cost of $5,000, program 2 is the less expensive alternative, no matter the discount rate. Since all funds are given up in the first year for program 1, these funds could not have been invested in alternative programs, while under program 2, $4,000 could accrue benefits due to investment in alternatives, making program 2 less costly in the long run.

Year	Program 1 cost	Program 2 cost
1	5,000	1,000
2	0	1,000
3	0	1,000
4	0	1,000
5	0	1,000
Present Value (3% discount rate)	$5,000	$4,717
Present Value (5% discount rate)	$5,000	$4,546
Present Value (10% discount rate)	$5,000	$4,170

Amortization or annualization. Amortization or annualization is a process which

distributes the total costs of a resource over its useful life. While some short-term

resources incur recurrent annual costs throughout the intervention or program and do not

need to be amortized, other resources are longer lasting (over a year) and can be used

beyond the duration of the intervention or program and must be amortized. Like the

discount factor formula, the annualization factor formula (Levin, 1983) accounts for the

interest rate (r) and the lifetime of the resource (n):

$$a\,(r,n) = \frac{[r(1+r)^n]}{[(1+r)^n - 1]} \tag{2}$$

This formula also considers depreciation and opportunity costs, or forgone interest on the

undepreciated portion of the resource, distributing the resource's forgone interest over its

lifetime to arrive at an equivalent annual cost (Levin, 1983). Figure 2 offers an example

of how costs are annualized.

Figure 2. Annualization example

District administrators purchased a computer server for their school network with a replacement cost of $2,000, a useful life of 5 years, and an interest rate of 3%. Using the annualization formula, the annualization factor is calculated to be 0.2183. Multiplying the annualization factor by the replacement cost of the computer yields an annual cost of $437. Hence, over 5 years, the computer costs a total of $2,185.

Year	Annual cost
1	437
2	437
3	437
4	437
5	437
Total cost (3% interest rate)	$2,184

In summary, since amortization and discounting serve different purposes, both

may be performed in a cost analysis. For the present study, amortization obtains the

equivalent annual costs of capital resources such as TV/VCRs and Second Step kits, and

discounting computes the net present value of the intervention across three years of

implementation. The processes of discounting and amortization are standard processes

used in cost, cost-benefit, and cost-effectiveness studies.

Cost, Cost-Benefit, and Cost-Effectiveness Studies

Ross, Barkaoui, and Scott (2007) identified 103 cost studies in education,

suggesting that these types of studies are popular. They reviewed a wide range of studies,

including studies that analyzed costs at elementary, secondary, and postsecondary school

levels; cost, cost-benefit, cost-effectiveness, and cost-utility studies; and studies that did

not focus on single preventative educational programs. Thus, many of their studies did

not meet the criteria for inclusion in this review, which required that studies use cost,

cost-benefit, and cost-effectiveness methodology to examine preventative and early

intervention educational programs that target student achievement and behavior outcomes in elementary school. A review of some of these studies follows, organized by type of intervention: early childhood services, class size reduction, one-on-one tutoring, whole-school reform, intervention teams, and behavior support.

   Early childhood services. Cost-benefit analyses of programs such as the California Abecedarian program and the High/Scope Perry Preschool program, which had outcomes studied through randomized experiments, have been performed to determine the ratio of costs to benefits of programs designed to provide quality preschool services (Barnett, 1985; Barnett, 1993). The California Abecedarian program was developed to provide intensive center-based preschool services ten hours a day, five days a week, to 112 children at-risk for low academic achievement.

   Masse and Barnett (2002) evaluated the cost of the High/Scope Perry Preschool program using a range of discount rates. Employing the ingredients approach to determine costs and estimating effects in monetary terms, Masse and Barnett (2002) found the total cost of the program at a public school setting to be $41,916 per student at a three percent discount rate, $40,427 per student at a five percent discount rate, and $39,041 per student at a seven percent discount rate.

   Although costs were high, Masse and Barnett (2002) claimed that the benefits of the Abecedarian program outweighed the costs. Using a three percent discount rate, they determined the program group special education placement rate (31%) to be lower than that of the control group (39%). This lower special education placement rate contributed to an $8,836 reduction in the costs of schooling per student (Masse & Barnett, 2002). Masse and Barnett determined benefits of the program by calculating monetary values of

benefits such as the earnings of the participants in the program, earnings of the participants' future children, maternal earnings, and costs saved from reduction of administrative costs due to a decrease in the number of Aid to Families with Dependent Children cases. The researchers also converted increased longevity due to decreased smoking into a benefit by attaching a monetary value to a year of life, and this cost and costs above were totaled to arrive at total program benefits in monetary terms. At a three percent discount rate, the benefits of the Abecedarian program were about four times the cost when the intervention was delivered in a child care setting and three times the cost if the intervention were to be replicated in a public school setting (estimated using market prices).

Masse and Barnett (2002) claimed that IQ gains made by the program group are durable, but they contradict themselves when they offer no explanation for the "fade" in IQ gains of these students by age 15. It appears that some early intervention programs demonstrate a trend in increasing IQ scores only throughout the duration of the program or for a few years. Once the program or its elements are removed, the effects on students' IQ scores fade, with scores eventually decreasing to the level of control students. Masse and Barnett's study was also based on only one IQ measure at each point in time, a different IQ measure each time, so the study would have benefited from the use of multiple measures—and identical ones, if possible—to assess intelligence.

In addition, although Masse and Barnett (2002) used random assignment, the study's small sample size and subsequently low power suggest that differences in outcomes between the control and treatment groups are subject to a great deal of error. They also used data that were not statistically significant; for example, their claim of a

reduction in the rate of smoking for the program group is based on nonsignificant results. Finally, attrition is a concern in Masse and Barnett's study: Some participants withdrew completely during the study while others reappeared later on. Since reported data include the total number of participants at each time period rather than all subjects who had been randomly assigned to treatment and control groups, it is difficult to determine how attrition may have affected outcomes. Although there are several limitations in Masse and Barnett's study, small sample size and attrition are plausible design flaws that weaken the credibility of their study.

The High/Scope Perry Preschool program also used random assignment to investigate the effects of preschool and home visits on 123 disadvantaged African-American children's educational achievement, economic success, and criminal activity (Schweinhart, 2003). In comparison to control children, children who participated in the High/Scope Perry Preschool program had higher literacy, school achievement, intellectual and language test scores at ages 19, 14, and 7, respectively. Cost-benefit was calculated using the same procedures as Masse and Barnett (2002): the discounted average cost of the program was $14,716 in 2003 and the program benefits were $105,324 per participant, or a cost-benefit ratio of 7.16 to 1 (Barnett, 1993; Schweinhart, 2003). Even at discount rates of 7% or higher, there were positive benefit to cost ratios (Barnett, 1993). Benefits were estimated by examining reductions in such things as costs of court settlements for would-be crime victims, reductions in justice system costs, reductions in welfare costs, and increases in taxes paid by the participants. Effect sizes were claimed to be moderate to large. Another benefit of the High/Scope Perry Preschool

program was its savings of $7,488 in school costs attributed to a reduction in special education services (Schweinhart, 2003).

The overarching benefits claimed for the Perry Preschool program are hard to believe, because the study suffers from a number of weaknesses, such as its questionable random assignment. For example, students with younger siblings were not randomly assigned to control and treatment groups but to the same group, and two students whose mothers did not fully participate in the program were moved over to the control group. The inadequate random assignment suggests that the supposed cost savings due to decreases in special education services and criminal activity and benefits in the form of increased student earnings may be due not to the program. Other program variables such as small staff to student ratios and strong parental support may have produced program outcomes.

Class size reduction. Another type of preventative or early intervention program is class size reduction (CSR). Most studies that investigate the cost-benefit of CSR are nonexperimental and claim unreliable cost estimates (Brewer, Krop and Reichardt, 1999). Brewer et al.'s cost estimates (1999) suggest that a national class size reduction down to 20 students in grades one to three cost $2.1 billion in 1998-1999. The general cost of class size reduction also depended on the specific class size reduction policy. For example, under a *flexible* policy, in which new classes are created only when the target class size is exceeded by 50% of the original class size, costs of national class size reduction were $4.7 billion in 1998-1999 and are projected to be $5.6 billion nationally in 2007-2008. Under a targeted policy, in which class size reduction is implemented only in schools where at least 50% of the student population receive free and reduced lunch,

costs of national class size reduction were $1.8 billion in 1998-1999 and are projected to be $2.1 billion nationally in 2007-2008. The costs of class size reduction also differs depending on teacher experience, so placing a cost on this intervention was challenging, as Brewer et al. (1999) pointed out.

Another class size reduction study, Project STAR, a well-known randomized trial on class size reduction from 22 to 15 students in Tennessee (47% more classrooms), learned that students assigned to smaller classes attained better test scores at a high program cost (Krueger, 2003). This experiment was ambitious in its amount of class size reduction, for most state policies do not support a class size reduction to below 20 students. It is expensive, costing $7,660 per student for 2.3 years (discounted at 3%) to reduce Tennessee's statewide class size in grades K to three to 15 students. Krueger (2003) indicated that the additional annual cost of serving a student in a smaller class was $3,501 in 1997-1998.

Apart from a .20 standard deviation increase in average test scores, benefits of Project STAR included a boost in the students' earnings and productivity later in life. Krueger (2003) concluded that at a four percent discount rate, benefits were $2 for each $1 spent. Thus, students' earnings later in life were double the costs of reducing class size from 22 to 15 students in Kindergarten to third grade. Krueger noted, however, that inner city schools may have been oversampled in Project STAR, and since minority students benefited more from smaller classes, the experiment may have overestimated benefits for students in general.

Another key flaw of Project STAR is its high attrition rate: only 48% of the students in the original experiment participated for all four years in the study (Hanushek,

1999), and attrition occurred annually. Finally, results cannot be generalized to other populations and settings, because the treatment was conducted in one state with a specific population not characteristic of the national population. It is possible that attrition threatens the internal validity of Krueger's (2003) study, for the treatment and control groups may become less equivalent than when they were initially randomly assigned.

One-on-one tutoring. Reading Recovery is another research-based early intervention whose cost-effectiveness has been controversial (Grossen, Coulter, & Ruggles, 1997; Marina & Gilman, 2003; Wasik & Slavin, 1993). A one-on-one tutoring program in reading for non-retained and non-special education first grade students who scored in the lowest 20% of their class in reading, Reading Recovery required certified teachers to design individualized programs for students. Students received 30 minutes of daily tutoring for a period of 12 to 20 weeks to help them augment their reading achievement to the class average (Marina & Gilman, 2003).

Wasik and Slavin's (1993) "best evidence" synthesis found inconsistent results for Reading Recovery. Although effect sizes ranged from medium to large, negative effects were detected on one test, Dictation, at the end of the program. Wasik and Slavin also indicated that a smaller percentage of Reading Recovery students (22%) than control students (31%) were either retained or assigned to special education. However, they asserted that by third grade, these differences in student retention and special education rates were no longer observed. It appears to be common that the positive effect of early educational programs on areas such as retention and special education rates, student achievement and IQ scores are not durable (Jones, Gottfredson, & Gottfredson, 1997; Masse & Barnett, 2002). More importantly, Wasik and Slavin's "best evidence" design,

advocated by Slavin, is less rigorous than the traditional meta-analysis, since it includes qualitative studies that do not report effect sizes. Specifically, Slavin's "best evidence" synthesis fails to establish clear and precise criteria for the selection of studies for inclusion. Also, without standardized measurement across selected studies, which is a serious flaw, Wasik and Slavin's best evidence synthesis may make inappropriate comparisons across studies.

Grossen et al. (1997) claimed that the program was expensive: 30 hours of Reading Recovery instruction cost $4,625 to $8,333 per student. In addition, most students required supplemental reading services after completing the program. Wasik and Slavin (1993) concurred with Grossen et al. that Reading Recovery had high costs: it cost $125,000 a year ($100,000 for two teacher salaries and $25,000 for training) for a school of 600 students.

Aside from its high costs, Reading Recovery proponents such as Marina and Gilman (2003) also exaggerated claims about its benefits (Grossen et al., 1997). For example, students who did not make progress in the program were not counted as participants, so outcomes only represented successes. Grossen et al. (1997) also observed that Metropolitan Achievement test scores failed to improve under Reading Recovery and contended that the program did not raise achievement levels overall. Grossen et al. made several valid points about the Reading Recovery cost-benefit study: it inappropriate reported achievement levels only for children who made progress in the program, it failed to raise student reading achievement in comparison to non-treatment students on standardized tests, and its excessive program costs outweigh the meager benefits to the students. While effect sizes for the first cohort of students were large at the end of the

implementation year (ES = 0.72), effects faded at one-year follow-up (ES = .29), as reported by Wasik and Slavin (1993).

In contrast to Grossen et al.'s (1997) critique of Reading Recovery, Marina and Gilman's (2003) recent cost-benefit evaluation of Reading Recovery suggested that Reading Recovery is much less expensive than traditional programs and produces positive outcomes. While Reading Recovery costs ranged from an undiscounted estimate of $2,300 to $3,500 per student, costs were much less than those of retention and special education placement (Marina & Gilman, 2003). According to the researchers, benefits were observed, for students obtained equal or higher reading test scores than average students by the end of the first year of the program (Marina & Gilman, 2003; Wasik & Slavin, 1993).

The Marina and Gilman study had a major limitation. The 51 students in the study all successfully completed the Reading Recovery program, so students who did not make progress were not included in calculations, as Grossen et al. (1997) emphasized. Because the performance of successful students may have inflated the estimated effects of the program, Marina and Gilman's results are misleading and overestimate the benefits of Reading Recovery.

Whole-school reform. A whole-school reform program is a process of using research-based approaches to improve the performance of an entire school by focusing on changes in all elements of the school. Several whole school reforms have been subjected to cost analyses. Although Success for All (SFA) incorporates one-on-one tutoring for disadvantaged students, it can be categorized as a whole-school reform as well, for the intent is to improve reading achievement overall at an SFA school (Madden, Slavin,

Karweit, Dolan, & Wasik, 1993). Reading tutors are certified teachers who design individualized, daily, 20-minute tutoring sessions and conduct 8-week reading assessments for their students. Students in the early grades are also regrouped during reading instruction based on their reading levels. A quasi-experimental evaluation of SFA suggested its benefits: students outperformed their matched control counterparts on tests after participating in the program, special education referrals and placements for learning disabilities diminished by about half in five schools, and retentions decreased in all schools (Madden et al., 1993). Because students were merely matched and not randomly assigned, however, the two groups of students are not equivalent with respect to true scores for groups, a key flaw in the study. As Jones, Gottfredson, and Gottfredson (1997) pointed out, "Whenever the school match is imperfect, errors of judgment that are difficult to document or detect and regression artifact in the student-level match may influence results" (p. 646). Thus, other factors apart from the intervention may have contributed to the appearance of positive results, suggesting that benefits may be exaggerated.

While Madden et al. (1993) found that costs of SFA ranged widely from $182,000 to $626,000 per school, close to King's (1994) independent estimates of $261,060 to $646,500, the savings from special educational testing and placements ($80,000 per year per school) and retentions ($263,000 per year per school) were great. In addition, Madden et al. (1993) made an argument about the possibility of reallocating schools' Title I funds to implement SFA, suggesting that schools would unlikely assume all costs. This assertion fails to take into account opportunity costs. Reallocation of costs would be an opportunity cost to schools, since fewer funds would be used to support students who

normally benefit from Title I academic services if these funds were funneled to SFA services.

A more recent evaluation of SFA investigated long-term cost-effectiveness for the students in the original SFA schools, as well as the students in quasi-experiment matched control schools (Borman & Hewes, 2002). The researchers found that the annual average cost of schooling through the eighth grade per low-achieving student in the SFA program was $72,800, based on the current market values for personnel salaries and training, materials and professional development. The cost of schooling through the eighth grade for matched control low-achieving students was $75,400. This difference suggests that the cost of schooling through the eighth grade per low-achieving SFA student was slightly lower than that of low-achieving matched control students. In their cost of schooling estimates, which included expenditures such as per pupil expenditures, costs of special education, costs of retention, and SFA costs, Borman and Hewes applied five percent discount rates to costs of retention for control students only. They contended that control schools' cost of retention should be discounted due to future costs incurred in retaining students instead of intervening in the present with SFA. They seem to ignore that SFA students were also retained, so discount rates should be equally applied to these costs of retention.

Borman and Hewes also claimed that benefits of the program outweighed its costs in years subsequent to the program. For example, SFA students attained and maintained higher CTBS/4 (eighth grade math and reading achievement test) scores than control students. Moreover, retentions were down for SFA students, and those in special

17

education spent one-half of the academic year in special education instead of three-fourths of the year, as compared to control students (Borman & Hewes, 2002).

Borman and Hewes' research design was seriously flawed from the outset: schools were inadequately matched, since SFA schools were some of Baltimore's lowest performing schools, while control schools were not categorized as low performing. Moreover, since schools were not randomly assigned and matching was inadequate, cost and effect comparisons between control and treatment students are unreliable. Finally, it is important to note that Slavin et al. (1992) and Borman and Hewes (2002) are developers of SFA or their colleagues or subcontractors, so their evaluations of SFA may present it in the best light. For example, an independent evaluation of SFA only found large positive effects of SFA in kindergarten, but not in grades one through three (Jones, Gottfredson, & Gottfredson, 1997). Moreover, Greenlee and Bruner's (2001) findings indicated that Title I schools without a SFA reading program outperformed Title I schools with SFA in place. Also, SFA led to negative effects on the reading achievement of students who were initially at or above grade level prior to SFA implementation.

While most of the previous studies described were cost-effectiveness or cost-benefit analyses, a cost analysis may also prove informative, if all programs compared were equivalent in their effectiveness in producing outcomes for students (King, 1994). In a cost analysis study of three whole-school reform programs, King (1994) compared SFA, Accelerated Schools and Comer's School Development Program (SDP). Focusing primarily on personnel costs, since these costs contributed most to total costs, King (1994) investigated additional budgetary expenses due to personnel, training, and staff development in the first few years the programs were implemented. In addition, marginal

societal costs, or parent and volunteer time, as well as opportunity costs, or time of existing personnel, were considered in the analysis. King (1994) also conducted a sensitivity analysis, which provided information about the lowest and highest cost extremes of each program. Her results suggested that SFA was the most costly program, with budgetary expenditures ranging from $261,060 to $646,500 per school, due to its requirements of additional staff, staff development and training, and a flat fee for program implementation. In addition, the program called for 1 to 2 hours a week of a reading tutor's time, staff time for advisory committee meetings, grade-level meetings, and parent workshops.

Although Traub (1999) concluded that SFA was an expensive program, finding the total average costs for all SFA models to be $140,000 per school, this cost does not fall within the range specified by King. Moreover, Traub classified the program's effectiveness as strong, meaning that it was evaluated by at least four rigorous outcome studies, though he neither identified the studies nor delineated the actual outcomes.

Perhaps due to its fewer staff requirements, Levin's (1987) Accelerated Schools program was much less expensive than SFA, ranging from a total of $45,000 to $116,000 (King, 1994). The bulk of the costs were attributed to training, staff development, and travel. Steering committee meetings, school cadres, extended day program monitoring and reflection time consumed up to 10 hours per week of staff time. Accelerated Schools also had the highest marginal societal costs, requiring up to 19 hours per week of additional staff and administrator time. Traub's (1999) estimate of a $130,000 per school average cost for all Accelerated School models was much higher. In addition, he classified the effectiveness of the program as marginal (evaluated by one rigorous study).

Comer's School Development Program (SDP) had similar costs to Accelerated Schools of $52,000 to $128,000 per school (King, 1994). Because costs were not disclosed by the program staff, King (1994) estimated these costs by using the extreme costs of the other two programs. She learned that SDP employed several staff members, had training costs, and a flat fee for initial implementation. Furthermore, given the variability in program implementation at schools, time estimates were difficult to make, but King determined that up to 10 hours a week were devoted to weekly team meetings, governance team meetings, extended day program, possible home visits, and parent participation. SDP had high marginal societal costs—teacher time—of up to 8.25 hours per week. In contrast to King (1994), Traub (1999) was able to estimate costs for SDP, albeit employing an unknown methodology, designating its costs as $130,000 per school on average for all variations of the program. In summary, King's (1994) study is notable for its detailed and comprehensive cost analysis.

Regarding effects of the SDP, Hayes, Comer, and Hamilton-Lee (1988) found positive effects with the Comer model in Michigan: their results revealed four-year test score gains for the program schools compared to overall district gains. However, the program schools were selected for participation based on their low levels of achievement and behavior and their principals' interest in participating in the program. Thus, the lack of random assignment to schools makes plausible the threats of regression to the mean and selection, suggesting unconvincing findings on the efficacy of the SDP.

Intervention teams. Although there have been several cost-effectiveness studies designed around various preventative and early intervention education programs, few studies have examined the cost-effectiveness of team-based interventions in schools.

Hartman and Fay (1996) compared the Instructional Support Teams (IST), a problem-solving team that offers instructional services to students with academic difficulties, with the special education process. To examine the effectiveness of IST, Hartman and Fay (1996) measured referrals to IST, referrals for psychological evaluations, the number of students placed in special education, and number of retentions. Cost measures included determining standard personnel salaries and benefits and surveying the average time that teachers and other school staff devoted to implementing IST. Since personnel were involved in both the special education process and the IST process, Hartman and Fay asked them to estimate their time spent in each process.

Basing costs on the salary and benefits and costs of time spent by school personnel involved in the IST and special education processes, costs were estimated at $670 per student for the IST process and $804 per student for the special education process. Due to increased personnel costs, IST cost $38,000 more per school than the special education process in the first year, when special education referrals and placements and retentions were included in cost estimates. However, over a ten-year period, IST cost $7,000 less per school at a five percent discount rate. In an elementary school of 500 students, the IST cost per student was $835 and the special education cost per student was $849 at a five percent discount rate. At a ten percent discount rate, IST and special education costs per student were virtually equivalent: the IST cost per student was $716 and the special education cost per student was $713. A ten percent discount rate seems unreasonably high: it makes more sense to select lower discount rates for programs and interventions based on lower expected returns on investments given current economic conditions. Hartman and Fay also use a zero percent discount rate in addition

to their five and ten percent rates; but, in this case, it would have been more appropriate to use a smaller range of discount rates (e.g. three to five percent).

Hartman and Fay's (1996) use of discount rates allowed for sensitivity analysis and ability to express the future costs of the programs using present values, but their range of discount rates was inappropriate. One recommended method for selecting discount rates is to start with initial discount rates of three percent and five percent and recalculating costs at higher discount rates (Levin & McEwan, 2001). If Hartman and Fay had employed a 3% discount rate, their results would have shown IST costs to be higher than those calculated at the 5% rate but lower than costs under the traditional approach. This information could have supported their argument that IST is less costly than the traditional approach.

Despite its costs, Hartman and Fay claimed that IST provided substantial benefits, especially in reducing the need for special education placements. Assuming that students spent about 10 years in special education, each reduction in the number of special education students meant a savings in 10 years of special education costs, or a $51,000 cost savings per student not placed in special education (Hartman & Fay, 1996). According to Hartman and Fay's (1996) sensitivity analysis, decreasing the number of years a student spent in special education was not nearly as effective as decreasing the number of students placed in special education.

IST did not reveal a substantial cost savings over the special education model. However, IST produced positive effects in terms of special education rates: in IST schools, the number of students placed in special education decreased from eight to seven students, or a 12.5 percent reduction in special education placements. The reduction in

placements enabled more students to be served through regular education, which offered more effective instructional services than special education at approximately the same costs, suggesting that IST was cost-effective. Hartman and Fay's conclusions are uncompelling, because they are based on comparisons of costs and effects between the schools prior to their IST implementation and schools that had been implementing IST, some of which were the same schools.

Behavior support. Scott and Barrett (2004) performed a case study cost analysis on Positive Behavior Support (PBS), another preventative team-based and school-wide intervention, which aims to prevent and reduce aggressive and impulsive behaviors. Under PBS, schools are expected to promote positive lifestyle changes for both students and school systems by addressing integrity in implementing the program and evaluating outcomes. To estimate costs, administrator salaries and staff and student time spent in disciplinary procedures and instruction were collected (Scott & Barrett, 2004). To determine the effectiveness of behavior strategies, only one measure, the reduction in number of student office discipline referrals and suspensions were counted. Results suggested that the frequency of behavior problems diminished across years of project implementation. In addition, Scott and Barrett (2004) claimed that a decrease in the amount of administrator time consumed by processing referrals and suspensions amounted to a cost savings of $6,024 per school in the first year and $6,932 in the second year.

Scott and Barrett's PBS study (2004) is limited in a number of ways, suggesting that results are questionable. First, they only included training costs in their analysis of the costs of PBS implementation. PBS requires a considerable amount of staff and

teacher time to design, implement, support, and evaluate interventions. The failure to include staff time in the cost analysis distorts the costs of the program, since staff time likely contributes to most of the costs. Second, the researchers attempt to calculate the cost savings of PBS by subtracting dollars saved in instructional and administrator time from costs of training. Since they have not taken into account the major costs of PBS—personnel time—their calculation underestimates PBS costs. Failure to include personnel time for an intervention that requires significant personnel time is a critical flaw that undermines Scott and Barrett's cost analysis. In addition, efficacy of PBS cannot be determined by one measure of the number of office discipline and suspension referrals, which are highly unreliable. Lastly, any decrease in problem behavior could be due to historic events related to implementing a new program in the school. According to PBS efficacy literature, a 90% reduction of problem behaviors in a given program is regarded as successful; based on this criterion, over half of the intervention outcomes met with success.

Foster, Jones, and the Conduct Problems Prevention Research Group (2006) evaluated the costs and effects of Fast Track, a multi-component, multi-grade program also aimed at preventing problem behavior, particularly aggression in at-risk, early-elementary-school-aged children. Using a randomized controlled trial, the researchers used cost-effectiveness ratios to compare the costs of avoiding conduct disorder, interpersonal violence, and criminal offenses between schools with and without Fast Track over a ten-year period. To compute cost-effectiveness ratios, they divided incremental costs of the program by incremental effect sizes. For example, they obtained an incremental program cost of $62,500 and an incremental effect size of .018 to

calculate the incremental cost-effectiveness ratio for avoidance of conduct disorder. This calculation resulted in a cost-effectiveness ratio of $3.48 million per case of conduct disorder averted, revealing that the Fast Track was not cost-effective. Foster et al. determined the maximum amount that society was willing to pay to avert cases of conduct disorder was $1 million per case, less than the cost-effectiveness ratio for avoidance of conduct disorder. The cost-effectiveness ratio for averting cases of conduct disorder was significantly higher than the cost-effectiveness ratios for averting interpersonal violence ($736,010 per act of interpersonal violence averted) and criminal offenses ($423,480 per crime averted). Foster et al. did not provide a breakdown of their methodology for calculating incremental costs and incremental effect sizes; thus, it is difficult to convince readers of the credibility of their cost-effectiveness ratios.

The cost and effect estimates were based on available and non-comprehensive budget information and limited outcome measures, and were all subject to much error. Estimating standard errors with the bootstrapping technique, Foster et al. (2006) admitted that standard errors for all three cost-effectiveness ratios were large. Another problem in the study was the use of only two measures to assess outcomes: (a) Diagnostic Interview Schedule for Children (DISC) to diagnose conduct disorder, and (b) Huizinga and Elliot's Self Report of Delinquency to document cases of interpersonal violence and criminal offenses.

On a much smaller scale than the behavior-related programs evaluated above, the Good Behavior Game (GBG), a universal-level elementary school-based preventative program, attempts to improve disruptive classroom behaviors and reduce future crime and substance abuse (Embry, 2002). Embry (2002) contends that the GBG is a simple

approach that rewards students' lack of negative behavior. Thus, it is low in resource cost while engendering strong, long-lasting positive effects. Based on a hypothetical implementation of GBG in the state of Wyoming, Embry claims that the GBG may cost an annual average of $200 per participant, or a total of $1,000,000 per year for 5,000 first and second grade students. Regarding benefits, Embry relies on state budgets and hypothetical projections to estimate that the savings from reductions in special education placements, correctional expenditures, and tobacco use may range from $15-20 million a year. Without providing an explanation of the methodology for his benefit projections, they seem inflated and unrealistic. Aos et al.'s (2004) independent cost evaluation of the GBG offers drastically different cost and benefit estimates. They indicate that GBG costs only $8 per participant, while producing $204 in societal benefits (a benefit-to-cost ratio of 25.92). Aos et al.'s cost estimates are understated, however. They chose not to include the costs of training for teachers delivering the intervention and teacher time spent implementing and monitoring the intervention.

In general, the cost evaluation studies concerning elementary school-based, preventative, behavior programs are inadequate due to research design and methodology problems; however, Aos et al.'s (2004) independent evaluation of a number of popular school-based preventative programs provides one of the better economic evaluation models. Aos et al. determined that only some of the programs analyzed offer positive benefit-to-cost ratios. For example, the Seattle Social Development Program (SSDP) is a structured school-wide intervention that aims to reduce risks and increase protective factors, focusing on students living in high crime areas. Teachers receive training in classroom management, parents in development of family and school relationships, and

elementary-aged students in school attitudes, behavior and academic achievement. Aos et al. found that the average cost per SSDP participant was $4,590, while the benefits resulting from a reduction in future crime totaled $14,426 per participant (a benefit-to-cost ratio of $3.14). Although Aos et al.'s evaluation offers better credibility than other cost analyses, it did not measure the key costs of personnel time. In particular, it did not estimate participant time spent in training, which were opportunity costs; hence, the costs of the SSDP are underestimated in the analysis.

     Summary of the literature. The literature regarding cost and cost-effectiveness of preventative school-based behavioral interventions is generally characterized by uncompelling research design and methods. Major concerns with previous efficacy research methods include failure to use rigorous approaches such as random assignment, proper matching of controls, and multiple outcome measures. On the cost side, limitations in research methods include exaggeration of benefits, failure to incorporate opportunity costs, and deficiencies in costing and selection of discount rates. The present study is part of a randomized trial that uses several outcome measures, and costs are assessed using the ingredients approach with sensitivity analyses. Therefore, the study addresses the key problems apparent in previous cost and cost-effectiveness research.

Cost-Effectiveness Analysis of Second Step

     The present study aims to investigate the cost-effectiveness of a preventative, school-based social competency program called Second Step, which strives to augment student's social skills to prevent problem behavior (Frey, Hirschstein, & Guzzo, 2000). Second Step is a social competency program in the form of a series of manual-based lessons directed by trained teachers. At the inception of the initial school year in which

the program is introduced, teachers typically participate in one-day training workshops and other school personnel participate in half-day workshops. Teachers deliver instruction in group settings in the skill areas of empathy, anger management, problem solving, and impulse control. Although instruction occurs once a week, teachers are encouraged to incorporate Second Step into daily instruction to allow students to apply the social skills taught (Frey et al., 2000).

Some published studies have examined the effectiveness of Second Step in decreasing problem behaviors or increasing student achievement (Frey, Nolen, Edstrom, & Hirschstein, 2005; Grossman, Neckerman, Koepsell, Liu, Asher, Beland, Frey & Rivara, 1997). Frey et al.'s (2005) trial of Second Step attempted to use teacher ratings, self-report and observations to assess the connection between students' social behavior and cognitions. They examined the effects of the intervention on social goals, attributions, satisfaction, and behavior, as well as the relationships among these variables. In the first year of intervention implementation, 47 second- and fourth-grade teachers were trained to deliver weekly Second Step lessons, and in the second year, 48 third- and fifth-grade teachers were trained. The mean percentage of intervention implementation was 79% in the first year and 83% in the second year, and most students did not receive the entire program.

Problems occurred at the onset of Frey et al.'s (2005) study. The control group consisted of specific classrooms within control schools that did not implement Second Step, rather than entire schools. For instance, classrooms that participated in the study received Second Step program materials and training, while the classrooms that did not participate in the study were considered control classrooms and did not receive these

perceived benefits. Hence, treatment diffusion may have occurred, because the control classroom participants may have been influenced by the intervention classroom participants in the same school. It is unknown whether or not the inability to implement the intervention throughout the entire school affected results suggesting the intervention group's improvement in behavior. Results indicated that the treatment group was more likely than the control group to demonstrate less aggression, more cooperation (among girls), more egalitarian satisfaction, and more prosocial goals. When divided into pairs, children who espoused prosocial goals and non-hostile attributions were more cooperative. Effect sizes were small (*ES*=.25) and effects were observed only among children who had initially exhibited high levels of antisocial behavior.

Although characterized as a randomized trial, Frey et al.'s assignment methods were flawed, for schools were not truly randomly assigned to intervention and control groups. In the year prior to the study, two-thirds of the recruited schools were randomly assigned to the intervention group and one-third to the control group; however, in the following year, four newly recruited schools were all assigned to the control group. Another limitation of the study is its high attrition rates: after two years, the attrition rate was 25.5% for the treatment group and 28.8% for the control group.

Grossman et al.'s (1997) randomized controlled trial evaluated the effects of Second Step on diminishing of student aggression and improving students' neutral and prosocial behavior. From their pre-test, post-test and six-month follow-up behavioral observations, Grossman et al. found that the second and third graders in the intervention group exhibited a reduction in overall negative behavior and physical aggression (the control group displayed an increase) and an increase in prosocial behavior relative to the

control group. Teacher- and parent-rating scales suggested no such positive effects, however. In addition, Grossman et al's small sample size of 12 schools and inclusion of only one year of intervention implementation likely contribute to the modest effects found on their rating scales. The inconsistency between effects found in the observations and ratings suggests that either of these assessment measures may be unreliable and may not accurately reflect the success of the intervention. In addition, Grossman et al. reported the significance of differences between the treatment and control group rather than effect sizes.

The few efficacy studies regarding Second Step only hint at the overall character of the literature on the efficacy of preventative school-based behavioral interventions—literature that often fails to use persuasive research design and methods. Rigorous research to support the effectiveness of Second Step and other commonly used intervention in schools is limited and only recently available; however, the "market" for high quality research is "evolving quickly" (Aos et al., 2004, p.1). The cost and cost-effectiveness literature on preventative interventions generally suffer from weak research design and methodology (see Appendix B for a summary table). Since society spends so much money on education, but there are few evaluations on both the efficacy and cost-effectiveness of school-based preventative interventions, research should focus on assessing the cost-effectiveness of these programs and interventions, including Second Step.

Despite the dearth of high quality efficacy and cost-effective research on Second Step, it is a popular program with widespread implementation; schools in all 50 of the United States and all Canadian provinces use the model (Committee for Children, 2006).

The format and content of Second Step is similar to other commonly used programs, and Second Step is supported by modestly positive results of efficacy studies. As a result, it was endorsed as an exemplary program by the Safe, Disciplined, and Drug-Free Schools Expert Panel (U.S. Department of Education, 2001). However, standards for efficacy trials are inconsistent among various panels and organizations. For example, the Society for Prevention Research (2004) published a list of rigorous criteria that must be met for a program to be categorized as efficacious or effective. Based on its criteria, Second Step would not be considered an efficacious program. However, because of its wide use and preliminary results, cost-effectiveness research on Second Step is warranted to make solid conclusions about its costs and effects as a school-based preventative intervention for elementary school students.

Research Questions

Cost, cost-benefit, and cost-effectiveness analyses concerning preventative and early intervention educational programs in elementary schools could in principle provide information for the scientific community to better inform administrators and teachers about how to maximize costs and ameliorate student outcomes. Unfortunately, few economic evaluations have examined elementary school-based, preventative, behavior programs. Scott and Barrett's (2004) and Foster et al.'s (2006) cost analyses are two examples of studies within the last fifteen years that employed a cost methodology to investigate the costs of an elementary school-based, preventative, behavior program. Thus, the paucity of cost research on behavior and related programs calls for further exploration of these programs. The efficacy research on which cost studies rely is also often weak.

The present study uses Levin and McEwan's (2001) cost-effectiveness methodology to compare the costs of Second Step and its effects on problem behaviors, prosocial behaviors and student achievement, to those of programs in control schools in a local school district within the context of a school-randomized trial.

Therefore, this study poses the following research questions:

1. What are the additional costs of implementing Second Step at treatment schools compared to those of usual and customary social and character development services at control schools?

2. What is the cost-effectiveness of an enhanced whole-school social competency intervention compared to the control set of programs in terms of self-reported aggression, academic grades, and MSA scores?

In answering these questions, this study seeks to produce information that will enable the scientific community to assist school administrators and officials in understanding the cost-effectiveness of an enhanced whole-school social competency intervention. Subsequently, the school district may be able to make informed policy decisions regarding whether or not to start the intervention in additional schools.

Methods

Introduction

A cost-effectiveness analysis evaluates program alternatives that share a similar

goal, to compare their costs and effects. Currently, both the control and treatment schools

implement a variety of formal and informal preventative interventions and programs that

target social and character development. Hence, since treatment schools already

implement other social and character development programs apart from Second Step, the

present study focuses on any additional costs required to implement an enhanced whole-

school intervention. The complex of social and character development programs at

control schools is considered one "control program."

In order for the enhanced whole-school social competency intervention to be cost-

effective, it should be more effective than the control set of programs in decreasing

problem behavior and increasing prosocial behavior if it involves incremental costs.

Cost data from the initial and second years (training costs), 2004-2005 and 2005-

2006, and the third year of Second Step implementation in the school district, 2006-2007,

are utilized.

The Outcome Study

The present study occurred in concert with an ongoing, larger, three-year

randomized controlled trial that investigated the efficacy of Second Step in 12 schools in

Anne Arundel County, in Maryland. After principals agreed to random assignment,

researchers selected 12 elementary schools that had never implemented Second Step to

participate. To account for differences in school size, ethnic composition, free-and-

reduced-meals (FARMS) percentage, and test scores, schools were matched based on

these characteristics. Within each pair of schools, researchers randomly assigned one school to the treatment group and one to the control group.

All students from first through fifth grade and all first- through fifth-grade teachers in the 12 schools were recruited to participate. From an initial student population size of 4,794 and a teacher population size of 210, the researchers initially recruited and received consent from 209 teachers and parents of 3,167 first- through fifth-grade elementary school students. Data from main outcome measures include student self-report surveys, teacher and parent ratings, report cards, and Maryland State Assessment (MSA) scores.

Intervention

A school-based social competency program, Second Step is a universal, problem behavior prevention curriculum that seeks to develop students' prosocial behaviors while preventing and reducing their aggressive and impulsive behaviors. It encompasses four skill areas of empathy, problem solving, impulse control and anger management, which are taught by trained teachers in the form of 30 to 35 minute, manual-based group lessons once a week. Lesson plans differ by grade in both content and number: grades one, four and five have 22 lessons, grade two has 17 lessons, and grade three has 15 lessons. Each treatment classroom received a Second Step kit that consisted of manuals, videos and activity sheets.

Prior to intervention implementation, the Committee for Children—the creators of Second Step—delivered one full day on-site workshop to all teachers, and a half day workshop to support staff at each of the six treatment schools. Make up trainings were available for those staff who were unable to attend the scheduled training dates. Trainings

34

focused on the details and principles of the program, expected implementation, and generalization outside of the classroom.

Progress monitoring of Second Step occurred through character development teams. Only treatment schools' teams monitored Second Step progress. The goal of these teams was to promote development of student character by monitoring all programs and activities related to character development. Since Second Step was such a program, the character development teams at the treatment schools were responsible for monitoring this program by maintaining contact with research staff. While the composition of these teams varied by school, individuals who participated on the team were usually also involved in Second Step activities (e.g. administrators, school counselors and teachers).

Integrity of program implementation was an important component of this particular Second Step implementation, and teachers were expected to complete implementation logs to show if they completed the steps required for each lesson. If teachers failed to complete a log, then the lesson was considered incomplete. At the end of each unit, research staff members collected and analyzed all logs and provided monthly feedback to the school counselors, who reported feedback to the character development team. Two types of feedback regarding each school involved in Second Step were supplied to the counselor: (a) percentage of lessons completed, and (b) percentage of key elements delivered. These implementation logs also enabled research staff to assess the amount of time devoted to curriculum delivery, which provided information about costs of personnel time. It is important to note that the Second Step program does not require such detailed progress monitoring: the level of progress

monitoring in this particular implementation of Second Step is no doubt much more thorough than typical implementations of the program.

Planning, implementation, support and supervision of Second Step required significant personnel time. Teachers were primarily responsible for preparation, implementation, and generalization activities related to Second Step. Specific activities in which teachers participated to implement Second Step are classified into the following seven categories: reading stories, showing video clips, facilitating group discussion, modeling role play, engaging students in role play, debriefing after role plays, and distributing handouts. Although teachers delivered Second Step lessons at specific times every week, they were encouraged to transfer learning by incorporating lesson principles into daily instruction through modeling, practicing, reinforcement and incidental learning (termed "transfer of training"). For example, at the end of the day, teachers asked students to "remember the day" and share with the class what they learned during the lesson.

Counselors contributed through teacher support, material distribution, troubleshooting, meeting facilitation, and counseling. They also operated the implementation monitoring system in the school, distributing and collecting logs and discussing the quantitative implementation data with individual teachers and the faculty as a group. Principals spent time planning and supervising Second Step activities. For two days in the summer after the first year of implementation, teachers, principals and counselors attended research staff meetings to discuss feedback on implementation, data collection concerns, and goal-setting. In the case of counselors, such meetings recurred monthly. Support staff members attended a mandatory half-day training about Second

Step to familiarize themselves with the program. The coordinator of guidance was the main district-level administrator for Second Step and coordinated the program activities. In addition, the project manager and data clerk were responsible for keeping records; organizing, making, and distributing materials; collecting logs and surveys; meeting with school personnel; and supporting counselors and school personnel. The coordinator of guidance, project manager, and data clerk are considered "program staff." Central office administrators, area directors, and other support and administrative personnel also spent time in project administration, management, and technical assistance.

Although the bulk of personnel time was spent in Second Step program implementation by school system personnel, members of the research team also spent some time in implementation. The term "research staff" includes the principal investigator, graduate research assistants, and the undergraduate research assistant. Specifically, the research staff involved in Second Step evaluation spent time preparing implementation feedback reports and attending meetings with school personnel.

Cost Analysis

This study employs the cost-effectiveness methodology described by Levin and McEwan (2001). It compares the social and character development programs in the control schools with the enhanced Second Step program in the treatment schools for differences in cost and cost-effectiveness. A cost-effectiveness analysis of the Second Step program without enhanced research elements is likely to reveal different costs and fewer effects. However, some of the research elements, specifically implementation measurement and feedback, are part of the enhanced intervention. To improve program implementation, several research elements were built into the implementation of Second

Step: clear implementation goals, consistent implementation measurement, and feedback about teachers' performance in implementing Second Step (Gottfredson, 2005). Any observed intervention effects are augmented due to these enhancements. Thus, in the cost analysis, certain research elements are treated as intervention costs. To aid the reader, Table 2 outlines the cost and outcome variables and explains how each are measured or collected.

Table 2

Cost and Outcome Variables

Variables	Category	Measures or Data Collection Methods
self-reported aggression	outcome	student self-report survey
academic achievement	outcome	report card academic grades
academic achievement	outcome	Maryland State Assessment (MSA) scores
personnel time (principal and counselor)	cost	personnel time use questionnaire—assessed the number of hours spent on specific character education activities per typical week
personnel time (research and program staff)	cost	personnel time use questionnaire—assessed the number of hours spent on activities related to program implementation during the third year
personnel time (teacher)	cost	implementation logs—measured percentage of time spent in curriculum delivery and provided contemporaneous estimates of time use
personnel time (teacher)	cost	transfer of training checklist—estimated the frequency of engaging in "transfer of training" outside formal Second Step instruction time
personnel time (teacher)	cost	supplemental teacher questionnaire—asked about the frequency and amount of time teachers devoted to formal character-education instruction

personnel salaries and benefits	cost	acquired from district contacts—obtained for all staff members who devoted time to treatment and control programs
facilities (classrooms, offices, conference rooms)	cost	acquired from district contacts and rental agreements—average market values of control and treatment spaces utilized for social skills programs are approximated from national median classroom cost per square foot (annualized)
equipment (TV/VCR player)	cost	acquired from district contacts—replacement cost of TV/VCR players (annualized)
materials (Second Step kits, poster sets)	cost	acquired from district contacts and accounting expenditures—replacement cost of Second Step kits (annualized)
transportation	cost	acquired from district contacts and accounting expenditures—mileage and travel distance required for trainers to travel to trainings and staff to travel to research team meetings
training (space, materials and supplies, trainers)	cost	acquired from attendance sheets, district contacts, and accounting expenditures—space cost approximated from cost of renting similar space in the county; personnel time determined from attendance sheets; materials and supplies and trainer costs estimated from accounting expenditures (initial training costs annualized)

Due to random assignment, treatment and control schools were equivalent in expectation with respect to all variables at the time of random assignment, including costs and programs implemented. Information about specific programs was revealed through a formal survey administered to teachers and principals shortly after the implementation of Second Step began. This survey asked about various social and character development programs in both control and treatment schools. According to results of the survey, the treatment and control schools were comparable in terms of number of and content of programs shortly after the implementation of Second Step began (see Appendix A for

summary table). One of the popular programs implemented was "Core Essentials," a program sponsored by fast food vendor Chick-fil-A. This program focused on highlighting specific character traits each month, which were reinforced through posters, charts, newsletters and coupons for free kid meals. Another popular program was called "Character Counts," which provided instruction and reinforcement for specific character traits through announcements, character conversations, newsletters, and videos and lessons in the classroom. Students also earned tickets and incentives for demonstrating appropriate character traits. All of the schools also incorporated counselor-directed, anti-bullying lessons into the guidance curriculum. Although Second Step was added to treatment schools and may have altered the nature of the schools' discipline activities, both treatment and control schools continued to implement their existing social and character development programs.

This study employs the ingredients approach (Levin & McEwan, 2001), a standard method of assessing costs of a target intervention. In this approach, average marginal costs of the enhanced whole-school social competency intervention in comparison to average marginal costs of the control program are estimated. Through working with district contacts and acquiring accounting and budget records, an itemized list of as many of the resources necessary to implement Second Step and other programs in the treatment school and the control programs as can be identified and their respective costs are compiled. The marginal costs of the social and character development programs are computed by adding the values of each cost component. The ingredients approach is described more fully in the following paragraphs.

Ingredients approach. The ingredients approach requires an estimate of all categories of costs involved in implementing the treatment and control programs in the treatment schools: personnel; facilities, materials and equipment; transportation; and other (which encompasses training and implementation costs). The researcher collaborated with district administrators to obtain information regarding the above costs, including school personnel salaries and benefits, personnel time and first-year and follow-up training and implementation costs. Mean personnel salaries and benefits by position for the district are calculated, although salaries and benefits vary within positions due to skill and experience levels. Hence, sensitivity analyses incorporate a range of salaries and benefits by position to estimate variability in total costs due to variability in salary and benefits. Costs are estimated for all school staff members who devoted any time to the treatment and control group programs, such as general education teachers, school counselors, and administrators. In addition, research costs include the costs of personnel time spent in record keeping; making, ordering and distributing materials; attending trainings, collecting surveys and implementation logs; meetings; and support and problem solving during implementation shortfalls.

Due to the heavy instructional component of school programs, personnel time contributes most to program costs. Personnel time for individuals involved in social and character development activities (such as Second Step) was determined by asking counselors and principals to complete a personnel time use questionnaire, which assessed the number of hours spent on social and character development activities per typical week. This questionnaire was developed by the investigators of the larger program evaluation. Investigators created separate questionnaires for counselors and principals, to

account for the difference in activities in which they engaged. The data instrument was modified in spring 2007 for the present study (Appendix C).

The personnel time use questionnaire included three parts. First, an explanation of the reasons for collecting time-use information was included at the top of the questionnaire. In addition, participants were informed about the voluntary and anonymous nature of the questionnaire and provided with information about returning their questionnaires. Next, the participant was requested to estimate retroactively the total number of hours spent in all social and character development activities per typical week. Third, each counselor and principal was asked to retroactively estimate the amount of time (in hours) spent in specific social and character development activities, distributed across the three domains of school-wide social and character development activity, classroom social and character instructional programs, and all other work activities. Participants who did not spend any time in a particular activity recorded a zero in the cell corresponding to the activity. In the total column, participants were expected to total the number of hours spent across the three domains to arrive at the total number of hours spent per activity per week.

Program and research staff were also asked to complete questionnaires about their time use. In the program staff time use questionnaire, program staff were asked to consider the totality of their work-related activities—divided into "Research," "Second Step Implementation," "Implementation of Other Social and Character Education Activities," "Other," and "Control" categories (see Appendix C)—and asked to estimate the total amount of time per year they spent in these activities. The "Research" category included activities such as record-keeping, distribution of forms, obtaining consent,

obtaining administrative buy-in for data collection, collecting data, and preparing

incentives for getting research surveys returned. In contrast, the "Second Step

Implementation" category described activities pertaining to program implementation,

such as ordering and distributing Second Step kits, collecting implementation logs,

meeting with counselors regarding Second Step feedback, meeting with school personnel

to devise strategies for encouraging staff participation in meetings and trainings and

getting implementation logs returned, and providing training support and attending

trainings for Second Step. For the "Implementation of Other Social and Character

Education Activities" category, program staff recorded time spent in the implementation

of social and character development programs outside of Second Step. The "Other"

category enabled program staff to estimate time spent in treatment schools performing

duties not covered by the first three categories. Finally, for the "Control" category,

program staff approximated time spent engaging in activities for control schools. To

guide them in approximating their time on the questionnaire, program staff received a list

of examples of research and implementation activities (see Appendix D). By contrast, in

a different, tailored time use questionnaire, research staff were asked to only approximate

the total amount of time per year spent in Second Step Implementation activities.

Teachers at Second Step schools completed implementation logs (see Appendix E)

for each lesson and unit of the Second Step program, which provided contemporaneous

estimates of time use in program implementation. In addition, these teachers provided

information on the frequency of engaging in "transfer of training" activities outside the

formal Second Step instruction time (see Appendix F). Counselors estimated the amount

of time typically consumed by each of these activities, based on their experience working

with teachers. Both control and treatment teachers also completed a supplemental teacher questionnaire in spring 2007, which asked about the frequency and amount of time teachers devoted to formal social and character development instruction.

District-level personnel, school budgets, and accounting records of project expenditures supply information on the cost of facilities, materials and equipment. Facilities, equipment and materials costs are counted in this cost analysis, whether or not they are existing resources, because some existing resources were reallocated for utilization in the implementation of Second Step. Since the facilities required in implementing Second Step typically consisted of classrooms, offices, and conference rooms that were jointly utilized by existing school functions and activities, costs of only those portions of facilities which were utilized specifically for Second Step are estimated. For facilities costs, the average market value of the control and treatment school spaces utilized for Second Step activities are approximated based on the national average elementary school classroom size and national median classroom cost per square foot. These costs are annualized to prorate the costs by time used.

Next, equipment costs consist simply of at least one TV/VCR player per grade level per school, shared among classrooms at the same grade level, which was needed to play VHS tapes that are included with Second Step kits. Since this equipment may have already existed in the schools and may be used for several years, the replacement cost of the TV/VCR player is multiplied by its annualization factor and percentage of time used to calculate its annual cost. Last, material costs include Second Step kits (photo-lesson cards, unit cards, VHS tapes, teacher and administrator guides, and handouts masters for homework) and poster sets. There are 18 sets of grade one through five kits for all the

treatment schools, plus additional kits for specific grades: two grade five kits, three grade one kits, and three grade two kits. Assuming that these kits are long-lasting, the cost of these kits is amortized (distributing the costs of an item over its lifetime) over a period of five years.

The third cost category, transportation costs, are minimal and encompass the mileage and travel distance required for Second Step trainers to travel to one full day and one half day Second Step workshops in the first year and refresher trainings in subsequent years, all of which took place within the county. Travel for teachers, counselors, administrators, and research team staff to research team meetings are also transportation costs. All of these transportation costs are estimated by obtaining accounting records of Second Step expenditures from district personnel and finding out district travel reimbursement rates and travel distance. While unreimbursed travel costs also contribute to total transportation costs, these costs are not included due to difficulty acquiring data about these costs.

Finally, the last cost category of other costs includes training and startup costs, which are collected from district administrators, accounting records, and budget information. The bulk of the training and startup costs are incurred in the first year of implementation of Second Step (2004-2005), so these data are collected retroactively through accounting records and amortized. Like facilities, materials and equipment, initial first-year training costs are capital costs that are amortized over five years. This amortization period assumes that teachers will benefit from the initial training for five years before requiring complete retraining. Refresher trainings are meant to review the initial training and are considered annual recurrent costs; hence, the costs of these

trainings are not amortized. Training costs for new staff consist of any costs related to Second Step training, such as costs of training space, training materials and supplies, trainers, and refresher trainings.

The cost of training space is estimated by determining the cost of renting similar spaces in Anne Arundel County. In addition, attendance sheets collected from the initial and refresher trainings reveal the number of people who attended trainings and the amount of time spent at trainings for all three years. Accounting records provide information on materials and supplies and costs of trainers. Similar procedures are followed to approximate personnel, materials and supplies, and facilities costs for the refresher trainings.

All of the cost categories described above are combined to arrive at average marginal costs of the treatment intervention. Average marginal costs are discounted at three and five percent discount rates, to estimate the future costs of the intervention if implemented in six additional schools for a three-year period.

Aside from personnel, facilities, materials, equipment, and transportation costs, opportunity costs must be considered in any cost analysis approach. The entire Second Step program is an opportunity cost, as student time that may have been spent in alternative academic instruction is reallocated to learning social skills under this particular program. In the experimental program, each grade in each treatment school set apart a block of time specifically for Second Step instruction. Although the total amount of instructional time remained unchanged, instructional time was reallocated to experimental program instruction. This time was labeled "Second Step" time and typically took the place of Media time or Specials time, to prevent students from missing

regular academic instruction. However, it was possible that teachers utilized Media and Specials time to provide supplementary math and reading instruction, so students may have been receiving less math and reading instruction if they missed Media and Specials time. Thus, the loss of students' learning time due to participation in Second Step is an opportunity cost that should be noted. Due to the potential loss of math and reading learning time, students may not have had as much time to learn the academic knowledge assessed on the Maryland State Assessment (MSA), and in turn, they may have been less likely to perform as well on the MSA. Therefore, in advance of knowledge about students' MSA scores, it is anticipated that this opportunity cost may be reflected in the reduction of students' scores on the MSA.

The average marginal costs of the control program are estimated using the ingredients approach and costs are organized into the categories of personnel; facilities, materials and equipment; transportation; and other. The counselors and principals who participated in the control program were requested to retroactively estimate the number of hours spent on social and character development activities per typical week. Due to the complexity of attempting to monitor implementation of a large set of other programs in the treatment and control schools, implementation logs and transfer of training logs were not completed at all by control teachers; and no implementation logs were used for programs other than Second Step. Control and treatment teachers filled out a supplementary teacher questionnaire that asked about time use in social and character development instruction.

Next, facilities, equipment and material costs of the control program are calculated in the same manner as those for the treatment program. Facilities included

offices, classrooms, and conference rooms, all of which were jointly shared with other

school activities. The control set of programs included one program called "Character

Counts," which required a TV/VCR player. Materials and supplies were essential and

may have been shared with other school activities. The annualization factor given in

equation 2 is used to compute the annual marginal costs of facilities and equipment. Thus,

the initial value of the facility or equipment is multiplied by its annualization factor to

estimate its annual costs.

Third, interviews with counselors at control schools suggested that only one

control school incurred transportation costs, since most activities related to the control

programs occurred within the schools and did not require participant and trainer travel.

The Positive Behavioral Intervention and Supports (PBIS) program, implemented at one

control school, required PBIS team members to attend at annual off-site trainings in the

summer, fall, and spring. Another program at another control school, "Kids Helping Kids

Campaign," encouraged students to travel to the community to participate in a service

activity: students and staff simply walked to a nearby site and did not require

transportation. School budgets and contributions by businesses or parent organizations

are reviewed to assess any other travel-related costs.

The last category of other costs encompasses startup implementation and training

costs, and these costs are collected by collaborating with school counselors. While many

of the control programs did not entail training, some such as the "Character Education

Program" and the "Bully Lessons and Anti-Harassment" provided in-service trainings for

staff. Since counselors directed these trainings, the personnel time use questionnaire (see

Appendix C) attempted to collect information on time spent providing trainings.

Effectiveness Analysis

With respect to effectiveness, effect sizes are estimated as part of the program evaluation of the experimental program. The main outcome measures that were examined in the study consist of the following: (a) spring 2007 student self-report aggression, (b) 2006-07 school year report card academic grades, and (c) spring 2007 Maryland State Assessment (MSA) scores for third and fourth grades. Effect size estimates of these outcomes are obtained from evaluation reports. While the efficacy study investigated several outcomes, only a limited range of outcomes are selected for the present study. Therefore, the self-reported survey scale of aggression is included. It is also possible that Second Step improved student achievement outcomes by allowing teachers to spend more time in classroom instruction than in behavior management. As a result, students may have had more of an opportunity to learn concepts on which they were graded and tested. Thus, report card academic grades and MSA scores are included. The following outcomes are excluded: teacher and parent ratings of social competency, other student self-report survey scales, and report card social scores.

First, the student self-report survey included 83 Likert items within 11 scales, and was completed by students in third and fourth grades in year three. This study examines only one of the survey scales that relate most to the goals of Second Step: Aggression. Survey data were collected in fall 2004, and the survey was re-administered in spring 2005, 2006 and 2007. In spring 2007, 72 percent of all third- and fourth-grade students completed the survey.

Next, the effect of the enhanced whole-school social competency intervention on student achievement is measured by comparing treatment and control students' state-

mandated standardized achievement test scores and student report card academic grades. In the outcome evaluation, third- and fourth-grade students' report card academic grades for the fourth quarter in spring 2004 (baseline year) and spring 2007 (third year) were averaged separately for each grade level to attain respective academic achievement composite scores. Higher composite scores indicated higher academic achievement.

Last, the achievement test is the Maryland State Assessment (MSA), administered to students in third and fourth grades. The MSA is a multiple-choice and short-essay test that assesses students in math and reading content, and it is administered annually to students in third through eighth grades. MSA scores reveal how students fare in learning the math and reading content established by the state curriculum, as well as how their performance compares to nationwide norms in math and reading. Third- and fourth-grade students' spring 2007 math and reading MSA scores were collected and averaged by grade to obtain composite MSA scores. Like report card composite scores, higher MSA composite scores mean higher achievement.

The effect of an enhanced whole-school social competency intervention on MSA scores is linked to the goals of Second Step. Second Step aims to increase prosocial behavior and decrease problem behavior, and in doing so, it may enable teachers to spend more time in academic instruction instead of classroom behavior management. If classroom time is redirected towards learning, students have more of an opportunity to acquire academic concepts tested on the MSA and will likely show improvement in test performance.

To calculate cost-effectiveness from the data listed above, the following formula is used (Levin & McEwan, 2001):

$$\text{Cost-effectiveness} = \text{Cost}_{\text{intervention}} / \text{Outcome}_{\text{intervention}} \tag{3}$$

The denominator is expressed in standard deviation units. These cost effectiveness ratios reveal the cost effectiveness of the enhanced whole-school social competency intervention in treatment schools as compared to the control programs in control schools.

Data Analysis

Regarding effectiveness, evaluation reports provide effect size estimates as well as standard error of the mean estimates and confidence intervals (CI) for the point estimates. Under the larger program evaluation, MSA scores and report card academic grades were analyzed to produce composite MSA and academic achievement scores. MSA scores on reading and math for the 2006-07 school year were averaged to obtain a composite MSA score, and scores of all students in third and fourth grades were standardized using *z*-scores. School year report card academic grades for 2006-07 were also analyzed in the same manner to generate a standardized academic achievement composite score. Thus, cost-effectiveness is represented as cost per standard deviation unit gained or reduced on the MSA and cost per standard deviation unit gained or reduced on academic grades.

Regarding cost, for personnel estimates in the ingredients approach, the time use of all Second Step and control program counselors and principals is aggregated by school position to obtain a mean and range on hours spent in social and character development activities. For teachers in Second Step schools, the average total time spent in implementation of Second Step instructional activities is provided by teacher time estimates. The year three implementation logs requested teachers to estimate the total amount of time spent on each Second Step lesson. An average and range are calculated

for these time estimates by calculating the average time for each teacher and then obtaining the grand mean and range. Similarly, average transfer of training time is calculated by multiplying activity time by frequency of activity for each lesson and averaging these amounts for each teacher before obtaining the grand mean and range. Contemporaneous time estimates of the total amount of time spent in Second Step activities are calculated for treatment teachers by summing average implementation and transfer of training times for each teacher and calculating the grand mean and range. Since comparable data for teachers are not collected in control schools, these implementation time estimates are utilized solely for assessing the marginal personnel costs of an enhanced whole-school social competency intervention and are not compared to costs of the control set of programs. For both control and treatment teachers, the retroactive estimate of the amount of time spent in social and character development activities per week is multiplied by the number of times engaged in these activities to obtain the average total amount of time spent in these activities per week. Retroactive estimates are used to estimate control and treatment teacher costs.

Time and salary and benefits of each position are analyzed to calculate personnel costs. The average cost (in dollars) per position per hour is determined by obtaining all school personnel salaries and benefits by position and dividing salary and benefits amounts by the number of hours worked in a typical school year, based on the Anne Arundel County Public Schools salary scales of 195 teacher work days and 260 administrator work days. For example, for teacher cost, (salary + benefits)/1560 hours per year = average cost per teacher per hour. Next, this cost is multiplied by the total number of hours spent on social and character activities per school position to calculate the

annual marginal cost of each position's time devoted to these activities. Finally, the costs of all individuals involved in the enhanced whole-school social competency intervention are totaled to obtain the total marginal treatment program personnel cost per six schools. Similar procedures are employed to determine control program personnel costs.

The next cost category is facilities, equipment, and materials and is estimated using annualization factors from Levin's annualization factors table (Levin & McEwan, 2001, Table 4.1). First, an annualization factor is applied to equipment cost estimates to compute annual costs of VCR/TVs, assuming that VCR/TV technology in schools has a lifetime of five years and depreciation rates of 3%, 5%, and 7%. The total cost of the Second Step kits—the main materials in the enhanced whole-school intervention—is collected from the district contact. Assuming that these kits will last at least five years before they are replaced, costs are amortized over a period of five years at 3%, 5%, and 7% depreciation rates, using the annualization factors table. The table is based on annualization formula 2.

Annualization recognizes that the value of a resource is consumed at the same rate from year to year during its lifetime, so it distributes the resource's forgone interest over its lifetime to arrive at an equivalent annual cost. To obtain the annual cost of the Second Step kits, the annualization factor is applied to the initial purchase price of the kits.

Under the transportation cost category, costs of traveling to training sites and other locales related to program implementation by trainers and implementation and research team staff is summed to arrive at the average marginal transportation cost for both control and treatment schools.

For the Other category, initial and refresher training costs for personnel time are computed by multiplying the average cost per personnel per hour by the number of hours spent in training and the number of personnel attending trainings. Initial training and startup costs are amortized. All cost categories (personnel; facilities, materials and equipment; transportation; and other) are totaled across both control and treatment schools to obtain the annual marginal cost of each program per student. Since the estimation of costs requires many assumptions, there is much room for error, so sensitivity analyses are conducted.

According to the research team, district contacts, and interviews with counselors at the treatment schools, only one program called "Bigs and Littles" necessitated volunteers: teachers volunteered to mentor low-income students one hour per week. Despite the minimum number of volunteers, their costs contribute to personnel costs and must be calculated. Average costs of volunteer time are calculated by assuming that the volunteers delivered services similar to those of teacher assistants (Levin, 1983). Thus, they should be compensated at the same hourly rate. Average hourly costs of teacher assistant time, based on annual salary and benefits, are multiplied by the total number of hours volunteered and the total number of volunteers. This figure is added to the marginal personnel cost of the treatment schools. In contrast, interviews with counselors at control schools suggested that there were no volunteers who contributed to personnel costs.

Regarding cost-effectiveness, cost estimates are converted into average marginal cost per student to determine the average marginal costs of the Second Step intervention. First, the cost categories are totaled for both control and treatment schools. Then, the average marginal cost of the intervention is used to estimate future costs of the

54

intervention for additional schools after three years of implementation. These three-year costs are discounted and converted to annual costs. Then, the annual discounted costs are divided by the average annual number of students enrolled at the control and treatment schools, to obtain the average marginal cost per student for implementing both the control programs and the enhanced whole-school social competency intervention in additional schools.

Using an example where the average marginal cost per student for both the control programs and the enhanced whole-school social competency intervention is the same, the effect sizes reveal cost-effectiveness. If the effect size for reducing aggressive behavior is .20 and costs are equal, then the treatment intervention is more cost-effective than the control intervention.

From the treatment and control cost and effectiveness estimates, cost-effectiveness ratios are computed, substituting effect sizes for the denominators:

$$\text{Cost-effectiveness} = \text{Cost}_{\text{intervention}} / \text{ES}_{\text{self-report}} \tag{4}$$

$$\text{Cost-effectiveness} = \text{Cost}_{\text{intervention}} / \text{ES}_{\text{academic grades}} \tag{5}$$

$$\text{Cost-effectiveness} = \text{Cost}_{\text{intervention}} / \text{ES}_{\text{MSA scores}} \tag{6}$$

The cost of the intervention program is divided by the effect size (*ES*) for each outcome included in the study (self-report of aggression, academic grades, and MSA scores). Hence, several cost-effectiveness ratios are estimated. However, calculating cost-effectiveness ratios for each of these outcome measures means that total intervention costs are divided by partial measures of outcomes, creating ambiguous cost-effectiveness ratios. In this situation, Levin, Glass, and Meister (1987) recommend averaging cost-effectiveness ratios for the outcomes. When these outcomes are "weighted equally,"

(Levin et al., 1987, p. 68) cost-effectiveness ratios offer clearer comparisons between interventions.

Since the numerator and denominator units generally are not equivalent in cost-effectiveness ratios, the ratio units are expressed as average cost per standard deviation unit gained or reduced per student in the outcome. Due to error in both costs and effects, and hence, error in cost-effectiveness ratios, a range of values are included to consider variability in both the numerator and denominator. Variability in cost-effectiveness ratios for different outcomes is examined by employing sensitivity analyses to determine the upper and lower bounds for the ratios. Specifically, a ratio of costs discounted at 3% and 5% and 80% upper and lower confidence limits of effect estimates are used (where z=1.28). Both 50% (where z=0.67) and 80% confidence intervals (CI) are also used to explore variability in effects under cost estimates discounted at 3% and 5%, as seen in Table 3.

Table 3

Sensitivity Analyses for Cost-effectiveness Ratios

Variability in Parameters	Range of Cost-effectiveness Ratios	
	Low	High
3% and 5% discount rates for cost estimates	$\dfrac{\sum\limits_{t=1}^{t}\dfrac{C}{(1+.05)^{t-1}}}{ES}$	$\dfrac{\sum\limits_{t=1}^{t}\dfrac{C}{(1+.03)^{t-1}}}{ES}$
80% confidence interval for effect estimates	$\dfrac{Cost_{\text{int}ervention}}{\overline{ES}-(1.28)SE_{\overline{ES}}}$	$\dfrac{Cost_{\text{int}ervention}}{\overline{ES}+(1.28)SE_{\overline{ES}}}$
3% and 5% discount rates for cost estimates and 80% confidence interval for effect estimates	$\dfrac{\sum\limits_{t=1}^{t}\dfrac{C}{(1+.05)^{t-1}}}{\overline{ES}-(1.28)SE_{\overline{ES}}}$	$\dfrac{\sum\limits_{t=1}^{t}\dfrac{C}{(1+.03)^{t-1}}}{\overline{ES}+(1.28)SE_{\overline{ES}}}$

3% and 5% discount rates for cost estimates and 50% confidence interval for effect estimates	$\dfrac{\displaystyle\sum_{t=1}^{t}\dfrac{C}{(1+.05)^{t-1}}}{\overline{ES}-(0.67)SE_{\overline{ES}}}$	$\dfrac{\displaystyle\sum_{t=1}^{t}\dfrac{C}{(1+.03)^{t-1}}}{\overline{ES}+(0.67)SE_{\overline{ES}}}$

Lastly, in an attempt to account for inherent variability in cost-effectiveness estimates due to sampling and measurement errors in estimates of effects and cost, sensitivity analyses are conducted by changing the variables in the assumptions about cost that are most likely to influence cost estimates. Variability in personnel salaries by position impact marginal costs and subsequently average marginal cost per student estimates, so a range of salaries by position (high and low) are included in total cost calculations.

In addition, a range of interest and discount rates are used. Annualization of costs includes interest rates at 3%, 5%, and 7%, and these rates are substituted for r in the annualization factor formula:

$$a\,(.03,n) = \frac{[.03(1+.03)^{n}]}{[(1+.03)^{n}-1]} \tag{7}$$

$$a\,(.05,n) = \frac{[.05(1+.05)^{n}]}{[(1+.05)^{n}-1]} \tag{8}$$

$$a\,(.07,n) = \frac{[.07(1+.07)^{n}]}{[(1+.07)^{n}-1]} \tag{9}$$

Finally, future costs of the enhanced whole-school social competency intervention are discounted at three and five percent rates, as seen in the present value formula:

$$PV = \sum_{t=1}^{t}\frac{C}{(1+.03)^{t-1}} \tag{10}$$

$$PV = \sum_{t=1}^{t}\frac{C}{(1+.05)^{t-1}} \tag{11}$$

<div align="center">**Results**</div>

Effectiveness

Effect sizes for the six treatment schools are estimated as part of the program

evaluation of the experimental program and are drawn from outcome evaluation reports

(Gottfredson, Harak, Nebbergall, Nese, & Shaw, 2008). Overall results of the outcome

evaluation show that there are no significant positive effects observed across parent and

teacher social competency ratings, the student self-report survey, report card social scores,

report card academic grades, and MSA scores for students in first through fifth grades

after three years of intervention implementation. The present study focuses only on a few

of these outcomes. Table 4 displays data for each of the three outcome areas examined

(report card academic grades, MSA scores, and self-reported aggression) in the third year

of implementation for third- and fourth-grade students. As seen in the table, after three

years of implementing the enhanced whole-school intervention, there were no positive

effects on student behavior and achievement.

Table 4

Year 3 Student Outcomes for the Enhanced Whole-School Social Competency

Intervention in Year 3

Outcome/Grade Level	Effect Size *(ES)*	Standard Error *(SE)*
Self-reported aggression – grades 3 and 4	- [a]	-
Report card academic grades – grade 3	-0.20	0.14
Report card academic grades – grade 4	-0.06	0.17
Standardized test scores (MSAs) – grade 3	-0.21	0.15
Standardized test scores (MSAs) – grade 4	-0.22	0.16

Self-reported aggression. One of the goals of Second Step is to decrease students'

aggressive and impulsive behavior. On a self-report survey with 83 Likert-type items and

11 scales, the Aggression scale related most to this goal. In spring 2007, the survey was

administered to third- and fourth-grade students to determine their own assessment of

their social development after participation in the intervention. Results of the self-

reported aggression scale reveal that there was little between-school variance (intra-class

correlation or ICC = 0.01). Because schools did not differ with respect to mean self-

reported aggression, treatment effects were not calculated.

Report card academic grades. It is also possible that Second Step could improve

student achievement by enabling teachers to spend more time in instruction and less time

in behavior management, giving students more opportunities to learn. As a result,

treatment students' academic grades should be higher than those of control students. In

the outcome evaluation, for each grade level, academic items (e.g. math and reading) on

third- and fourth-grade report cards were averaged for the fourth quarter in spring 2007.

Results suggest that the enhanced whole-school intervention had a small negative effect

on academic grades for third-grade students (*ES*=-0.20) and virtually no effect on

academic grades for fourth-grade students (*ES*=-0.06).

Standardized test scores. A standardized achievement score, based on the

Maryland State Assessment (MSA) served as a second academic outcome. In the

outcome evaluation, third- and fourth-grade students' math and reading MSA scores were

averaged by grade level to obtain composite MSA scores. Results suggest that the

enhanced whole-school social competency intervention may have had a small negative

effect on MSA scores for both third-grade students (*ES*=-0.21) and fourth-grade students

(*ES*=-0.22), but this did not approach statistical significance. Had there been a significant difference in student MSA scores, that difference would have supported the speculation that an opportunity cost of implementing Second Step was lost instructional time.

Costs

As a school-wide intervention, Second Step is not simply an add-on program, for one of its intents is to replace existing discipline activities through its goal of preventing problem behavior at the school level. Prevention should ultimately decrease the amount of time that counselors and principals spend in discipline activities, enabling them to reallocate their time to other social and character development activities aimed at improving school-wide outcomes. By training teachers to provide modeling, instruction, and feedback to students on core skills in empathy, problem-solving, anger management and impulse control, Second Step may help students to change their attitudes and thoughts about social situations and deal with these situations prosocially. As a result, Second Step may reduce the need for discipline referrals. Thus, the amount of time consumed by counselors and principals in discipline management may be no greater or even less in treatment than in control schools. Consequently, a reduction in the reported amount of counselor and principal time spent in discipline and social and character development activities may be expected.

Personnel. Costs of the intervention consist of the salaries and benefits of all personnel involved in implementation of the intervention. Tables 5, 6, 7 and 8 display total time spent in social and character development activities by principals, counselors, teachers, and program and research staff. For each treatment and control principal and counselor, the total number of hours devoted to school-wide and classroom character

development activities are combined with each other to calculate the total time spent in these activities. Thus, principals' and counselors' total time encompass time allocated to character development teams, which aim to monitor the progress of social development programs.

There is little difference between treatment and control principals' and counselors' time spent in social and character development activity (Table 5). Table 5 displays initial principal time use estimates for both treatment and control groups, as well as adjusted time use estimates after excluding one extremely high principal estimate from each group. Adjusted time use estimates are necessary, because initial principal time estimates are probably unreliable. One treatment and one control principal reported extreme time estimates that elevated mean principal time estimates. These individuals both reported spending over 50 hours in social and character development activity, even though the typical work week only contains 40 hours. Although unlikely, they also both avowed devoting the majority of their time to social and character development activity and working a total of 70 or more hours a week. Thus, these two principals' time estimates are excluded in calculations of principal time to reflect more accurate average principal time use.

According to the counselor and principal time use questionnaire, treatment principals reported spending an average of 4.5 hours per week in social and character development activity, while control principals reported spending an average of 8.3 hours. On the other hand, treatment counselors reported spending an average of 16.8 hours per week in social and character development activity, while control counselors reported spending an average of 19.4 hours. As anticipated, treatment principals and counselors

actually spent slightly less time in social and character development activity, which

encompassed Second Step, than their control counterparts.

Table 5

Year 3 Principal and Counselor Time per Week (in Hours) Spent in Social and Character

Development Activity per Treatment and Control School

	Treatment Time			Control Time		
Position	Mean	*SD*	Range	Mean	*SD*	Range
Initial Principal (n=6 treatment, n=5 control)	15.2	26.2	56.0	18.0	23.0	66.5
Adjusted Principal (*n*=5 treatment, n=4 control)	4.5	2.4	6.0	8.3	8.5	18.0
Counselor (*n*=5 treatment, n=4 control)	16.8	6.8	14.5	19.4	2.5	6.0

Note. Initial principal time use estimates include two principals who provided extreme time use estimates. Adjusted principal time use estimates exclude these two principals.

Implementation logs and transfer of training checklists provide contemporaneous

estimates of the amount of time that treatment teachers spent in implementation of

Second Step. In years two and three, the implementation logs provided a space for

teachers to estimate the total amount of time that they spent delivering each Second Step

lesson. Table 6 presents the average, low, and high teacher time estimates for both formal

instruction and transfer of training activity. Treatment teachers devoted an average of

0.47 hours per week to classroom instruction in Second Step. Thus, most teachers

adhered to lesson delivery within the 0.50 hour per week allocated to Second Step

instruction. Even so, the range of time spent on each Second Step lesson was large: the

highest lesson completion time was 0.88 hour per week and the lowest was 0.25 hour per

week. Outside of formal Second Step instruction, teachers engaged in "transfer of

training" activities for an additional 0.53 hours per week on average, which was

comparable to the amount of time devoted to formal Second Step instruction. To compute

"transfer of training" time, the amount of time spent on each transfer of training item is multiplied by the frequency of engaging in each item. Then, these amounts are averaged for each individual and averaged across all treatment teachers. "Transfer of training" time ranged widely from a high of 1.76 hours per week to a low of 0.00 hours per week. The average total teacher implementation time for the enhanced Second Step intervention was 1.00 hours per week, with a high implementation time of 2.19 hours per week and a low implementation time of 0.35 hours per week.

In contrast, the supplemental teacher questionnaire provides retroactive time estimates for both treatment and control teachers; hence, these time estimates enable comparison between treatment and control groups and are used to estimate teacher costs. The frequency and amount of time devoted to formal social and character development activities are multiplied to determine control and treatment teacher time. Control teachers reported spending 0.28 hour per week in formal social and character development instruction, compared to 0.53 hour per week—almost twice the amount of time—by treatment teachers. Some control teachers reported that that they spent no time in formal social and character development instruction, while others reported spending over 1.33 hours per week. On the other hand, treatment teachers reported spending a range of 0.13 to 2.58 or more hours per week in formal social and character development instruction.

As can be seen from Tables 6 and 7, contemporaneous time estimates do not match retrospective time estimates for the treatment-school teachers. Implementation logs and transfer of training reports (contemporaneous estimates) reveal that treatment teachers spent a total of 1.00 hour per week in Second Step activities, while the supplemental teacher questionnaire (retrospective estimates) indicates that they spent

0.53 hour per week in formal social and character development instruction. These figures

differ by a factor of 1.89. It is possible that treatment teachers did not include transfer of

training time in their estimates of formal social and character development instruction, in

which case the average completion time of 0.47 from Table 6 is comparable to the

average total implementation time of 0.53 from Table 7.

To compare time estimates between the two methods, multiplying the

retrospective estimates by 1.89 yields 1.00 hour per week spent in formal social and

character development instruction by treatment teachers and 0.53 hour per week by

control teachers. In principle, both contemporaneous and retrospective time use estimates

should be equivalent, but the discordant estimates underscore the existence of error in

time estimates. Due to the ability to obtain comparable retrospective estimates of teacher

time from both treatment and control schools, retrospective time use estimates are used to

approximate teacher costs.

Table 6

Year 3 Contemporaneous Estimate of Teacher Time per Week (in Hours) Spent in Second

Step Implementation per Treatment School

Time-use Estimate	Mean
Completion of curriculum lessons (*n*=113)	
Average completion time for all teachers	0.47
Highest teacher completion time	0.88
Lowest teacher completion time	0.25
Transfer of training (*n*=110)	
Average transfer of training time	0.53
Highest transfer of training time	1.76
Lowest transfer of training time	0.00
Total implementation (*n*=113)	
Average total implementation time	1.00
Highest total implementation time	2.19
Lowest total implementation time	0.35

Table 7

Year 3 Retrospective Report of Teacher Time per Week (in Hours) spent in Formal Social and Character Development Instruction per Treatment and Control School

	Mean Total Implementation Time	
Time-use Estimate	Treatment (*n*=108)	Control (*n*=81)
Average total implementation time	0.53	0.28
Highest total implementation time	2.58	1.33
Lowest total implementation time	0.13	0.00

Program and research staff attempted to improve teacher implementation of the whole-school social competency intervention by instituting the following components: setting implementation standards, providing feedback on teacher performance, and removing obstacles to implementation. With these components in place, teachers would learn how to improve their implementation of the enhanced whole-school intervention, leading to an increased likelihood of positive program outcomes. According to the outcome evaluation data, overall implementation integrity of Second Step was high across all three years. On average, teachers exceeded the implementation goal of completing 80% or higher of lessons for the entire Second Step curriculum, set by the research team and school staff. Based on contemporaneous self-report, teachers implemented the majority of Second Step lessons. Specifically, in year one, teachers completed an average of 91% of the lessons (*n*=111); in year two, they completed an average of 96% of the lessons (*n*=109); and in year three, they completed an average of 91% of the lessons (*n*=113) (Gottfredson et al., 2008).

Table 8 shows the total amount of time program and research staff reported spending per year in the categories of "Second Step Implementation" and "Implementation of Other Social and Character Education Activities." The program

manager reported spending about 401 hours out of 1650 contracted work-hours, or about

24% of her time, on Second Step implementation. The data clerk reported spending no

time in implementation activities. The coordinator of guidance spent only five hours in

Second Step implementation, because most of her time was allocated to continued

management and administration of other social and character development activity. In

regards to the research staff, the principal investigator spent only four hours in program

implementation, while the three graduate research assistants and the undergraduate

research assistant spent a combined 186 hours and 142 hours, respectively, in program

implementation.

Table 8

Year 3 Total Program and Research Staff Time Spent in Implementation of Second Step

and Other Social and Character Development Activity (Outside of Second Step) per Six

Treatment Schools

	Total Amount of Time Spent per Year (Hours)	
Position	Second Step Implementation	Implementation of Other Social and Character Development Activity
Program Staff		
Program manager	401	300
Data clerk	0	0
Coordinator of guidance	5	9
Research Staff		
Principal investigator	4	--
Graduate research assistants total (*n*=3)	186	--
Undergraduate research assistant	142	--

Note. Based on an assumption of 1650 teacher work hours and 2080 administrator work hours per year for program staff.

Based on the average salaries and benefits for all treatment and control personnel

involved in implementation of social and character development programs, costs per hour

for each position are calculated (Table 9). Salaries and benefits are reported in 2006-2007

dollar amounts. For principals, counselors, and teachers, highest and lowest salaries and

benefits are shown to reflect the range of salaries and benefits. By chance, there is a

greater range in salaries and benefits for counselors at control schools, compared to their

treatment school counterparts. Mean costs per hour by position are comparable between

treatment and control schools. For program and research staff, salaries and benefits are

aggregated in Table 9 to protect confidentiality, but individual costs per hour are

multiplied by total amounts of time spent per year in implementation of social and

character development activities to calculate marginal costs of personnel. The costs per

hour used are the following: (a) one program manager at $52 per hour, (b) one data clerk

at $13 per hour, (c) one coordinator of guidance at $54 per hour, (d) one principal

investigator at $80 per hour, (e) three graduate research assistants, each at $13 per hour,

and (f) one undergraduate research assistant at $10 per hour.

Sensitivity analyses assess the consequences of variability in cost per hour and

annual cost for principals, counselors, and teachers, based on highest and lowest salaries

and benefits per position (Tables 9 and 10). Cost per hour varies considerably between

treatment and control counselors, suggesting a greater range of counselor experience

level at control schools due to chance.

Table 9

Year 3 Salaries and Benefits by Position per Six Treatment and Six Control Schools

Position	Treatment Schools Total Cost (*n=6*)	Control Schools Total Cost (*n=6*)
Principals (*n=6* control, *n=6* treatment)		
Average salaries and benefits	$126,070	$119,427
Highest salary and benefits	$131,191	$132,377
Lowest salary and benefits	$114,523	$105,665

Total salaries and benefits	$756,421	$716,562
Average cost per hour	$61	$57
High cost per hour	$63	$64
Low cost per hour	$55	$51
Counselors (*n*=6 control, *n*=7 treatment)		
Average salaries and benefits	$71,814	$81,563
Highest salary and benefits	$76,544	$111,889
Lowest salary and benefits	$63,552	$46,197
Total salaries and benefits	$502,698	$489,377
Average cost per hour	$46	$52
High cost per hour	$49	$72
Low cost per hour	$41	$30
Teachers (*n*=149 control, *n*=170 treatment)		
Average salaries and benefits	$62,591	$64,533
Highest salary and benefits	$70,292	$72,744
Lowest salary and benefits	$52,195	$56,574
Total salaries and benefits	$12,727,746	$11,747,852
Average cost per hour	$40	$41
High cost per hour	$45	$47
Low cost per hour	$33	$36
Volunteers (*n*=33)		
Cost per hour	$10	--
Program staff salaries and benefits	$250,128	
Cost per hour	$120	--
Research staff salaries and benefits	$236,831	
Cost per hour	$103	--

Note. Based on 195 teacher work days and 260 administrator work days. Program staff and research staff salaries and benefits are aggregated to protect confidentiality. Volunteers include teachers who participated in a mentoring program at one of the treatment schools, but do not include the parent volunteer who taught students sign language as part of a tolerance lesson.

Table 10 compares the annual cost of personnel's time spent in social and character development activities by treatment and control schools. The cost of personnel time is the product of the estimate of hourly cost and number of hours per year. Teacher costs in Table 10 are calculated from retroactive reports of time spent in formal social and character education instruction.

Generally, treatment schools incurred slightly less overall marginal personnel costs. This result is surprising, since the treatment intervention required additional personnel time from teachers and program and research staff, so total marginal treatment

personnel costs would be expected to be higher. The reduced amount of time treatment

principals and counselors spent in social and character development activities account for

the lower marginal treatment personnel costs. Costs of volunteer time contribute to

treatment school personnel costs. Volunteers include 33 teachers who participated in the

"Bigs and Littles" mentoring program one hour a week at one of the treatment schools.

The parent volunteer who taught students sign language for a tolerance and diversity

lesson at a treatment school is not included in this count, due to lack of information about

time spent by this individual. In year three, treatment schools spent a total of $400,210 on

personnel who devoted time to social and character development activities, while control

schools spent a total of $421,332. When looking at annual marginal costs of personnel's

time, there is a small difference in the negative direction ($21,122) between the six

treatment and six control school costs. This difference is highly sensitive to variability in

the range of cost estimates, particularly differences in salary and benefits. For example,

when using the lowest salary and benefits to estimate personnel costs, marginal personnel

costs for treatment schools were $52,480 higher than those of control schools.

Table 10

Year 3 Marginal Cost of Personnel's Time Spent in Implementation of Social and

Character Development Activities per Six Treatment and Six Control Schools

Position	Treatment	Control	Difference
Principals (*n*=4 control, *n*=5 treatment) sum	$85,644	$147,607	-$61,963
High cost	$88,452	$165,734	-$77,282
Low cost	$77,220	$132,070	-$54,850
Counselors (*n*=4 control, *n*=5 treatment) sum	$182,988	$231,192	-$48,204
High cost	$194,922	$320,112	-$125,190
Low cost	$163,098	$133,380	$29,718
Teachers (*n*=81 control, *n*=108 treatment) sum	$93,428	$42,533	$50,895
High cost	$105,107	$48,758	$56,349
Low cost	$77,078	$37,346	$39,732

Volunteers (*n*=33)	$12,870	--	--
Program manager	$20,852	--	--
Data clerk	$0	--	--
Coordinator of guidance	$270	--	--
Principal investigator	$320	--	--
Graduate research assistant (*n*=3)	$2,418	--	--
Undergraduate research assistant	$1,420	--	--
Total annual marginal costs	$400,210	$421,332	-$21,122
High costs	$426,361	$534,604	-$108,243
Low costs	$355,276	$302,796	$52,480

Note. Annual marginal high and low costs are based on the highest and lowest salaries and benefits of principals, counselors and teachers at the six treatment and six control schools. The first row for principals, counselors, and teachers is the sum of actual marginal costs for the six schools in each group, based on the total number of participants whether or not they responded to the time use questionnaires. Teacher costs are based on retroactive reports of time spent in formal character education instruction. For principals, counselors, and teachers, *n* refers to the number of individuals who responded to questionnaires about time use.

Facilities, equipment and materials. Total and annual marginal **c**osts of facilities used by treatment and control schools in 2007 dollars are displayed in Table 11. These costs are calculated based on the national recommended average elementary classroom size of 900 square feet (Tanner, 2000) and the national median cost per square foot of $152.05 (Abramson, 2008). Because each treatment and each control school teacher has his or her own classroom, there are 113 treatment classrooms and 95 control classrooms. Thus, the total classroom cost is $15,463,485 for six treatment schools and $13,000,275 for six control schools. However, costs of school facilities differ by the amount of time treatment and control teachers spent in delivering formal social and character development instruction. Table 11 uses retroactive self-report estimates of teacher time to assess marginal costs of facilities used. Treatment teachers reported spending about .53 hours per week in the classroom for Second Step implementation, or 1.3% of the 1560 contracted school hours, and control teachers reported spending about .28 hours per week in the classroom for implementation of control programs, or 0.7% of the contracted school hours.

A sensitivity analysis explores annual marginal facilities costs at 3%, 5%, and 7%

interest rates, assuming that schools have an effective lifetime of at least 25 years (The

David and Lucile Packard Foundation, 2007). Under a three percent interest rate,

treatment school facilities ($11,544) cost more annually than control school facilities

($5,226). Greater facilities costs for treatment schools are also observed under five and

seven percent interest rates.

Table 11

Marginal Costs of Facilities Used for Social Development Activities per Treatment and

Control School, Annualized at 3%, 5%, and 7% Rates

School	Marginal Costs (Average Market Value)	Annualized at 3%	Annualized at 5%	Annualized at 7%
Total per school				
Treatment	$33,504	$1,924	$2,377	$2,875
Control	$15,167	$871	$1,076	$1,301
Total per six schools				
Treatment	$201,025	$11,544	$14,263	$17,250
Control	$91,002	$5,226	$6,457	$7,809

Note. Based on 900 square feet classrooms and $152.05 per square foot; 1.3% facilities usage by treatment schools and 0.7% facilities usage by control schools; and an assumed lifetime of 25 years.

Equipment for treatment and control schools include TV/VCR players.

Replacement and annual marginal costs of equipment and materials are shown in Table

12. Each of the 12 treatment schools and two of the control schools uses one TV/VCR

player per grade level (grades one to five) for social and character development programs,

at a replacement cost of $403 per unit. For grades one to five, there are a total of 18

Second Step video lessons, which run about 10 minutes per lesson. Since there are about

three classes at each grade level, each treatment school used the TV/VCR players for a

total of 540 minutes, or nine hours per year. Assuming that the TV/VCR players are

71

generally used 2 hours per month or 20 hours per year on average, the TV/VCR players are utilized 45% of the time for Second Step instruction. Therefore, each treatment school spent $181 per grade on the TV/VCR player, for a treatment total TV/VCR cost of $5,441 ($n=30$). For the two control schools that use TV/VCR players for social and character development activities, six 25-minute videos were shown as part of a Character Counts program. These two control schools used the TV/VCR players for a total of 150 minutes, or two and a half hours per year. Assuming average general use of 20 hours per year, the TV/VCR players are utilized 12.5% of the time by control schools for social and character development activities. Hence, the control schools spent a total of $100 ($n=2$) on the TV/VCR players.

Materials for treatment and control schools consist of Second Step kits and poster sets and Positive Behavior Intervention Supports (PBIS) incentives for one control school. Treatment schools purchased 26 Second Step kits ($17,864) and 3 poster sets ($229). One control school implemented PBIS and spent $2,000 on incentives. A sensitivity analysis annualizes this equipment at 3%, 5%, and 7% interest rates over a five-year period. Treatment schools spent over ten times the amount spent by control schools on equipment and materials for social and character development programs.

Table 12

Yearly Annualized Costs of Equipment and Materials per Six Treatment and Six Control Schools, Annualized at 3%, 5%, and 7% Rates

Equipment and Materials	Total Initial Costs, Not Annualized	Yearly Costs Annualized at 3%	Yearly Costs Annualized at 5%	Yearly Costs Annualized at 7%
TV/VCR players				
Treatment ($n=30$)	$5,441	$1,188	$1,257	$1,327
Control ($n=2$)	$100	$22	$23	$24

Second Step kits (treatment only)	$17,864	$3,901	$4,126	$4,357
Second Step poster sets (treatment only)	$229	$50	$53	$56
PBIS incentives (control only)	$2,000	$437	$462	$488
Total costs				
Treatment	$23,534	$5,139	$5,436	$5,740
Control	$2,100	$459	$485	$512

Note. Based on $403 per TV/VCR player; 45% TV/VCR player usage by treatment schools and 2.5% TV/VCR player usage by control schools; and an assumed lifetime of 5 years.

Transportation. According to the school district grant accountant, travel by program staff to training and meeting sites accounted for the marginal third-year transportation cost of $1,852. Marginal transportation costs by position are presented in Table 13. Among research staff, graduate research assistants took 11 round trips to a central office site to meet with program staff for purposes of further program implementation and were reimbursed at a rate of .485 cents a mile. Thus, the cost of transportation for graduate research assistants in the third year was $320. In the first year, approximately $200 was spent for travel for one of the trainers who delivered character education and social skills training to staff.

The one control school that implemented PBIS incurred marginal transportation costs for travel to off-site locations for PBIS trainings. Four PBIS team members took a total of four round trips to off-site locations for two full-day summer trainings and two half-day fall and spring trainings, and they were reimbursed at .485 cents a mile. Although another control school social development program required travel outside of school, there were no transportation costs, because students traveled by foot to a nearby church to make lunches for a homeless shelter housed temporarily in the church. Therefore, in the third year, the marginal cost of transportation for the control schools was $369. Due to its extensive training and evaluation components, the enhanced whole-

school social competency intervention is more costly in transportation than the control set of programs.

Table 13

Year 3 Marginal Transportation Costs for Trainer, Program and Research Staff per Six Treatment and One Control Schools

Position	Total
Treatment	
Program staff (*n=3*)	$1,852
Graduate research assistants (*n=3*)	$320
Trainer (*n=1*)	$200
Control	
PBIS team – counselor, administrator, teachers (*n=4*)	$369
Total marginal transportation costs	
Treatment	$2,372
Control	$369

Other (training and startup). Total marginal training costs for treatment schools encompass the costs of personnel time, trainer fees, and facility rentals (see Table 14). In the first implementation year, staff members in treatment schools attended initial full-day trainings centered on the details, principles, and implementation of the enhanced whole-school intervention; thus, the bulk of the training costs were in the first year. It is assumed that staff will benefit from the training for at least five years, so first-year training costs are annualized. In the second implementation year, refresher trainings were conducted to review content covered in the first year's training. These trainings are recurrent costs and are not annualized. In the third implementation year, trainings included both refresher trainings and trainings for new staff who had not received the initial training. Hence, only new staff training costs are annualized. Third year training costs included one full-day workshop with a speaker for about 50 personnel. This workshop contributed to over half of the training costs in the third year, which explains

the higher third-year training costs. A sensitivity analysis explores treatment school

marginal training costs at 3%, 5%, and 7% interest rates over a five-year period.

For control schools, marginal training costs consisted of costs of personnel

attendance at two full-day and two half-day PBIS trainings every year during the course

of the character education grant. Each year, the PBIS trainings cost a total of $3,438, for a

three-year total of $10,314. Since all PBIS trainings were recurrent costs, these costs are

not annualized. In year three, treatment school marginal annualized training costs were

$25,844 more than control school training costs, under 3% interest rates.

Table 14

Marginal Training Costs per Six Treatment Schools, Annualized at 3%, 5%, and 7%

Rates

Training Year	Total Marginal	Annualized at 3%	Annualized at 5%	Annualized at 7%
Year 1 training costs	$88,928	$19,418	$20,540	$21,689
Year 2 training costs	$12,165	--	--	--
Total costs (refresher plus year 1 annualized cost)		$31,583	$32,705	$33,854
Year 3 training costs	$23,851	--	--	--
Refresher trainings	$5,969	--	--	--
Total minus refresher training	$17,882	$3,905	$4,130	$4,361
Total costs (refresher plus years 1 and 3 annualized costs)		$29,282	$30,639	$32,009
Total three-year training costs	$124,944	--	--	--
Refresher trainings	$18,134	--	--	--
Total minus refresher training	$106,810	$23,323	$24,670	$26,050

Total marginal costs. Based on an annualization rate of 3% for facilities,

equipment and materials, and trainings, the average annual marginal cost of all

ingredients required for the treatment intervention is $424,969 under a 5% discount rate,

compared to $410,634 for the control programs. Average annual marginal costs for the

six treatment schools are higher than those of control schools. However, there is a lower

annual marginal cost per student for treatment schools due to the greater number of

treatment students. Table 15 summarizes the annual marginal costs for each ingredient

category, as well as the annual marginal costs of all ingredients combined per six

treatment and six control schools.

Using a sensitivity analysis, marginal costs of all ingredients are discounted at 3%

and 5% to project variations in total marginal costs over a three-year implementation

period. Under a 3% discount rate, the treatment intervention will cost $1,299,198, while

the control programs will cost $1,255,193. Thus, in three years, the treatment intervention

will cost over $44,005 more than the control programs per six additional schools. Under a

5% discount rate, the treatment intervention will cost $1,274,907, while the control

programs will cost $1,231,903. Again, if implemented for three years, the treatment

intervention will cost $43,004 more than the control programs per six additional schools

with the 5% discount rate. Due to larger number of treatment students, the treatment

intervention will cost $24 less per year than the control programs under a 3% discount

rate and $23 less per year under a 5% discount rate.

To calculate cost per student estimates, discounted average annual marginal

program costs are divided by the average numbers of first- through fifth-grade students

enrolled each year. Changes in enrollment occurred throughout the course of the

interventions, due to typical student mobility in the school district; thus, not all students

received the entire Second Step or control interventions. Therefore, using the average

numbers of students enrolled each year considers typical student mobility and is more

realistic than using the total numbers of students who participated all three years or the

total numbers of students who participated in the third year. Across three years of

treatment and control interventions, the average number of first-grade through fifth-grade

treatment students enrolled is 2,511, and the average number of control students is 2,137.

Table 15

Total Marginal Costs of All Ingredients per Six Treatment and Six Control Schools,

Discounted at 3% and 5% Rates

Ingredient	Treatment Cost	Control Cost	Difference
Personnel			
Principals	$85,644	$147,607	-$61,963
Counselors	$182,988	$231,192	-$48,204
Teachers	$93,428	$42,533	$50,895
Volunteers	$12,870	$0	$12,870
Program staff	$21,122	$0	$21,122
Research staff	$4,158	$0	$4,158
Total annual costs	$400,210	$421,332	-$21,122
Facilities, equipment and materials (annualized at 3%)			
School space	$11,544	$5,226	$6,318
TV/VCR players	$1,188	$22	$1,166
Second Step kits/posters	$3,951	$0	$3,951
PBIS incentives	$0	$437	-$437
Total annual costs	$16,683	$5,685	$10,998
Transportation	$2,372	$369	$2,003
Training (annualized at 3%)			
Year 1	$19,418	$3,438	$15,980
Year 2	$31,584	$3,438	$28,146
Year 3	$29,282	$3,438	$25,844
Annual marginal costs of all ingredients			
Year 1	$438,683	$430,824	$7,859
Year 2	$450,848	$430,824	$20,024
Year 3	$448,547	$430,824	$17,723
Total three-year marginal costs			
discounted at 3%	$1,299,198	$1,255,193	$44,005
discounted at 5%	$1,274,907	$1,231,903	$43,004
Average annual marginal cost			
discounted at 3%	$433,066	$418,398	$14,668
discounted at 5%	$424,969	$410,634	$14,335
Annual marginal cost per student			
discounted at 3%	$172	$196	-$24
discounted at 5%	$169	$192	-$23

Note. Costs are calculated based on an average of 2,511 treatment students enrolled and 2,137 control students enrolled.

Although the difference between the marginal costs of treatment and control interventions is minimal, startup costs of the enhanced Second Step intervention are important to consider. There were more treatment students than control students, but there were also more treatment classrooms (*n*=113) than control classrooms (*n*=95). Thus, the average class size for both treatment and control schools was identical—an average of 22 students—so the startup costs of the treatment intervention do not depend on class size. If implemented in six additional elementary schools, startup costs for the first year of implementation of the enhanced Second Step intervention are $438,683.

Sensitivity analyses for costs. To assess the sensitivity of total marginal costs and annual marginal cost per student to cost assumptions, these costs are recalculated by varying interest rates and salaries and benefits (Tables 16 and 17). Total cost comparisons over a three-year period are somewhat sensitive to differences in interest rates and highly sensitive to differences in salaries and benefits. Differences between total marginal treatment and control costs are largest when using low salaries and benefits, followed by high salaries and benefits. The annual marginal treatment cost per student is lower than the control cost under all variations of interest rates and high salaries and benefits. Marginal training costs necessitated by the enhanced whole-school social competency intervention and reduced counselor and principal time spent in social and character development activities account for the main cost differences.

Table 16

Sensitivity Analysis for Total Marginal Costs per Six Treatment and Six Control Schools,

Discounted at 3% and 5% Rates Over Three Years

Total Marginal Costs	Treatment Cost	Control Cost
Total costs, annualized at 3%		
under a 3% discount rate	$1,299,198	$1,255,193
under a 5% discount rate	$1,274,907	$1,231,903
Total costs, annualized at 5%		
under a 3% discount rate	$1,311,466	$1,258,855
under a 5% discount rate	$1,286,944	$1,235,497
Total costs, annualized at 7%		
under a 3% discount rate	$1,324,620	$1,262,873
under a 5% discount rate	$1,299,849	$1,239,440
Total costs, high salaries and benefits		
under a 3% discount rate	$1,376,175	$1,585,207
under a 5% discount rate	$1,350,456	$1,555,794
Total costs, low salaries and benefits		
under a 3% discount rate	$1,169,071	$909,842
under a 5% discount rate	$1,147,195	$892,960

Table 17

Sensitivity Analysis for Annual Marginal Cost per Student, Discounted at 3% and 5%

Rates Over Three Years

Annual Marginal Cost per Student	Treatment Cost	Control Cost	Difference
Annual cost per student, annualized at 3%			
under a 3% discount rate	$172	$196	-$24
under a 5% discount rate	$169	$192	-$23
Annual cost per student, annualized at 5%			
under a 3% discount rate	$174	$196	-$22
under a 5% discount rate	$171	$193	-$22
Annual cost per student, annualized at 7%			
under a 3% discount rate	$176	$197	-$21
under a 5% discount rate	$173	$193	-$20
Annual cost per student, high salaries and benefits			
under a 3% discount rate	$183	$247	-$64
under a 5% discount rate	$179	$243	-$64
Annual cost per student, low salaries and benefits			
under a 3% discount rate	$155	$142	$13
under a 5% discount rate	$152	$139	$13

Cost-Effectiveness

To evaluate the cost-effectiveness of the enhanced whole-school social competency intervention, both costs and outcomes are combined to obtain cost-effectiveness ratios. The interest lies in the intervention's marginal cost-effectiveness, which is estimated with incremental cost-effectiveness ratios (ICERs). These ratios indicate that the enhanced whole-school social competency intervention is overall slightly less costly per student and no more effective than the control programs. Table 18 presents the values of the ICERs for the costs of the enhanced whole-school social competency intervention for the two student achievement outcomes. A treatment effect was not computed for self-reported aggression due to the lack of between-school variance. Thus, assuming an effect of zero, the ICER for that outcome cannot be calculated because the denominator would be zero. ICERs are computed by dividing the marginal cost per student (under a 5% discount rate and 3% interest rate) by the incremental effect of the program. In calculating ICERs, the marginal cost per student is a negative figure, to indicate that the treatment program costs slightly less per student than the control program.

An intervention's cost-effectiveness may differ for each outcome. For Second Step, total intervention costs are divided by partial measures of outcomes, resulting in ambiguous ICERs. In an attempt to address this problem, Levin et al. (1987) suggest averaging ICERs for alternative measures of similar outcomes. Table 18 reports the average of ICERs for both achievement outcomes. Ratio units express the average cost per standard deviation unit gained or reduced per student in each outcome.

Therefore, after three years of implementation, the enhanced whole-school social competency intervention will cost $69 less per student, and the best estimate is that no positive effects or small negative effects in student academic grades and MSA scores will be achieved. As depicted by the ICERs, the enhanced Second Step intervention is only somewhat less expensive than the control programs, and it is no better than the control programs at producing positive effects. Thus, it is no more cost-effective to implement the enhanced Second Step intervention than the control programs.

Table 18

Incremental Cost-effectiveness Ratios (ICERs) for Second Step per Standard Deviation (SD) Unit Gained or Reduced per Student

Outcome	Cost	Effectiveness	ICER
Self-report aggression	-$69	-[a]	-[a]
Academic grades	-$69	-0.26	$265
MSA scores	-$69	-0.43	$160
Average of ICERs			$213

Note. Costs are based on marginal costs per student under a 3% interest rate and 5% discount rate, and effectiveness is based on effect sizes for outcomes for third- and fourth-grade students in the third year of implementation.
[a] ICC = 0.01, treatment effects not calculated, so ICER cannot be calculated.

Sensitivity analyses for cost-effectiveness ratios. Tables 19 and 20 depict the sensitivity of achievement outcomes to variations in cost assumptions. Since the point estimates (ICERs) for academic grades and MSA scores are based on 5% discount rates, these estimates are equivalent to the lower bound ICERs of the first set of estimates in Tables 19 and 20. Point estimates fall between the absolute values of the lower and upper bound confidence intervals for effect estimates. For academic grades, ICERs fluctuate the most under differences in confidence intervals for effect estimates, especially with 50% confidence limits for effect estimates.

For MSA scores, ICERs vary greatly under different confidence intervals for effect estimates, especially with 80% confidence limits for effect estimates. Hence, the cost-effectiveness of each outcome is sensitive to particular variability in discount rates for costs and confidence intervals for effect estimates. In general, the enhanced whole-school social competency intervention costs slightly less to achieve no positive effects.

Table 19

Sensitivity Analyses for Cost-effectiveness Ratios for Academic Grades per Standard Deviation (SD) Unit Gained or Reduced per Student

	Range of Cost-effectiveness Ratios	
Variability in Parameters	Low	High
3% discount rate for high cost estimate and 5% discount rate for low cost estimate	$265	$277
80% CI for effect estimates	$105	-$504
3% discount rate for high cost estimate and 5% discount rate for low cost estimate Lower bound of 80% CI for low effect estimate and upper bound of 80% CI for high effect estimate	$105	-$526
3% discount rate for high cost estimate and 5% discount rate for low cost estimate Lower bound of 50% CI for low effect estimate and upper bound of 50% CI for high effect estimate	$148	$1,377

Table 20

Sensitivity Analyses for Cost-effectiveness Ratios for MSA Scores per Standard Deviation

(SD) Unit Gained or Reduced per Student

	Range of Cost-effectiveness Ratios	
Variability in Parameters	Low	High
3% discount rate for high cost estimate and 5% discount rate for low cost estimate	$160	$167
80% CI for effect estimates	$83	$2,078
3% discount rate for high cost estimate and 5% discount rate for low cost estimate Lower bound of 80% CI for low effect estimate and upper bound of 80% CI for high effect estimate	$83	$2,169
3% discount rate for high cost estimate and 5% discount rate for low cost estimate Lower bound of 50% CI for low effect estimate and upper bound of 50% CI for high effect estimate	$108	$324

Discussion

Although the cost and cost-effectiveness of preventative school-based behavioral interventions have been widely investigated, flaws in research design and methods limit the credibility of findings. Previous efficacy research suffers from inadequate random assignment and matching of controls, while cost research suffers from the lack of thorough cost assessment and gross exaggeration of benefits. Moreover, despite uncompelling research regarding its efficacy, Second Step is a popular intervention embraced by school systems across the United States and Canada. The present study attempts to address these flaws through its use of efficacy data from a randomized trial that incorporates several outcome measures and the ingredients approach with sensitivity analyses to assess costs.

Cost-effectiveness analyses are normally performed after the effectiveness evaluation of the intervention has been demonstrated (Rossi, Lipsey, & Freeman, 2004); thus, it is recognized that conducting the present study may be premature in light of the deficiency of previously conducted rigorous outcome evaluations. However, as mentioned earlier, there was an ongoing larger scale experimental study of an enhanced whole-school social competency intervention in the Anne Arundel County Public School System; therefore, the timing of the present study was fortuitous.

Results suggest that the enhanced Second Step intervention will cost $69 less per student than the control programs in three years (under a 5% discount rate), and it is no more efficacious than the control programs in improving student achievement. Cost differences between treatment and control schools occurred across all four cost categories. Treatment schools had lower personnel costs than control schools, due to a shorter

84

amount of treatment counselor and principal time spent in discipline and social and character development activities. For the facilities, materials, and equipment category, treatment schools sustained higher costs, due to the additional costs of Second Step kits and more frequent use of school space and equipment. Treatment schools also incurred higher costs for transportation, since several annual social and character development trainings and meetings required travel. Finally, treatment schools sustained higher training and startup costs due to the additional training costs required for implementation of the enhanced whole-school intervention. Although total three-year marginal costs for treatment schools are higher, this higher cost is due to the fact that there was a larger number of treatment than control students. The annual marginal cost per treatment student is lower than the marginal cost per control student.

Moreover, cost differences are observed between treatment and control personnel. A sensitivity analysis of personnel salaries and benefits shows variability in costs per hour between treatment and control counselors, suggesting differences in treatment and control counselor experience levels due to chance. Also, since social skills instruction is an essential component of the enhanced whole-school intervention, treatment teachers spent more time in formal social and character development activity than control teachers. It is plausible that results from retrospective self-reports are underestimates of treatment teachers' time use, for the results of the contemporaneous self-reports imply that teachers spent almost twice the amount of time in Second Step implementation. Despite high implementation integrity—teachers implemented over 90% of the lessons—no significant positive effects are observed in the three outcomes examined.

In summary, Second Step costs essentially the same as and is no more effective than the control set of programs after a three-year implementation period. Based on the evidence of efficacy and costs, it makes little difference whether the enhanced Second step intervention or the control set of programs is implemented. Hence, on the grounds of cost-effectiveness, this intervention cannot be preferred over the usual and customary social and character development programs in control schools.

Interpretation of Differences in Effects

Even though there is no incremental efficacy, the enhanced Second Step intervention may have been effective in comparison to not implementing any programs at all. A cost-effectiveness analysis compares costs and effects among programs that share similar goals; therefore, comparisons of efficacy are relative and could depend on which programs are compared. For example, if Second Step is compared with programs that are less efficacious, then Second Step would appear to be more efficacious. Alternatively, if a comparison is made between outcomes for Second Step schools and schools that implemented no social and character development programs, it is possible that Second Step schools may reveal more improvement in outcomes.

Interpretation of Differences in Costs

In the current study, the most unanticipated results are the lower costs of the treatment intervention. At the time of random assignment, the only difference between treatment and control schools was the addition of the enhanced Second Step intervention to treatment schools. Thus, one would expect costs of the treatment intervention to be more than those of the control intervention. In contrast to expectations, costs of the treatment intervention are actually less than the costs of the control intervention,

suggesting that the implementation of Second Step may have altered the nature of treatment school programs and practices.

Because treatment principals and counselors devoted less time to social and character development activities than their control counterparts, it is likely that the implementation of Second Step allowed them to redistribute their time to other activities. Second Step is not an add-on program, but rather a program that substitutes for an alternative opportunity that would have been implemented instead. As proposed earlier, Second Step may have changed existing discipline practices. It may have prevented problem behavior at the school level, consequently reducing the number of discipline referrals and the amount of time counselors and principals spent in discipline activities. These individuals' time may have been spent in alternative ways. For example, perhaps their time was reallocated to programs that promoted schoolwide academic achievement. Therefore, costs could be less for the treatment intervention if the amount of treatment principal and counselor time spent in social and character development activities was reduced and reallocated to other alternatives. Since principal cost is the largest cost difference between treatment and control schools, it is possible that an important outcome of the enhanced Second Step intervention is its modification of discipline practices and the subsequent effect on principal time use.

Limitations

The present study is limited in that it was carried out in one school district that underwent three years of intervention implementation. Due to the unavailability of cost data for all three years, third-year and retrospective training cost data are used to estimate total and future marginal costs. While third-year self-report personnel time estimates are

conservative because they do not reflect first-year startup and training activities, a cost assessment of three-year training costs captures personnel time spent in training. In addition, because this cost study is an add-on to an existing study, methods of estimating time spent in social and character development activity differed across school positions. For example, personnel time use questionnaires estimated counselor and principal time use in all schools, while implementation logs estimated teacher time use only in Second Step schools.

Also, the time use questionnaires may not have truly captured the principals' or the data clerk's actual time use. The initial mean treatment and control principals' time use estimates were elevated due to two principals' particularly high estimates, so these two estimates are removed from mean principal time use calculations. It is likely that other principals also encountered difficulty in making time use estimates. Some principals who did not return initial surveys were interviewed over the phone, and these principals had a difficult time ensuring that the total number of hours spent per typical week matched the total number of hours spent in all activities per typical week.

Whatever else may be the case, the estimate of principal and counselor time is based on a small number of principals, implying large standard errors, and is less precise than teacher estimates. There is no reliability estimate for time-use data for principals and counselors. Teacher estimates are most reliable, because they are based on both contemporaneous logging of activities and recall of amount and frequency of time spent in activities. Recording their time using two methods may also have increased the salience of tracking their time, presumably augmenting the accuracy of teacher time use estimates.

In addition, the data clerk may have misconstrued the time use questionnaire, for her self-report allocated all of her time in the third year to the research category, although research staff reported that she devoted some time to intervention implementation. Thus, her reported time is not included as part of the program staff implementation costs. As a consequence of potentially unreliable self-reports of time use, personnel cost estimates are sensitive not only to inherent variability in salary and benefits but also to error.

Another limitation is the potential threat of missing information about counselor and principal time expended on programs. Three counselors and one principal who participated all three years of the enhanced Second Step intervention did not respond to the third-year time-use questionnaire; hence, it is possible that the loss of their data could affect the average time use estimates for counselors and principals. There was a higher attrition rate for control counselors (17%) than for treatment counselors (8%). Similarly, there was a higher attrition rate for control principals (8%) than treatment principals (0%). There is no reason to believe that non-respondents had different characteristics than respondents. However, with no data to compare non-respondents and respondents, it is not possible to rule out differential attrition between the treatment and control groups as a threat to the validity of the time-use data.

The survey format for time use estimates from small samples is also a limitation, for surveys are known to be unreliable and small sample sizes imply large standard errors, as is also true for time diaries. However, compared to time diaries and observation, the survey format for estimating time use may lead to overestimates of daily and weekly time (in hours) spent on activities, so surveys are often regarded in the cost-effectiveness literature as the least reliable time measure (Robinson & Godbey, 1999). Moreover,

people may only recall days when the activities being surveyed were the most salient (Juster & Stafford, 1991). As a consequence, the days not recalled will not be represented on the survey, and this missing data may impair the accuracy of reports as an estimate of an individual's actual time use. Due to their limitations, surveys should only be used to assess the time use of frequent and consistent activities (Juster, 1985). The use of the time use and supplemental teacher questionnaires in the current study is acceptable, for they asked about activities that counselors, principals, and teachers performed regularly and consistently. The most reliable and valid methodology for measuring time use is observation throughout the duration of the activity in question (Juster, 1985), especially when coupled with simultaneous completion of time diaries (Robinson & Godbey, 1999). Implementation logs are time diaries that provide contemporaneous time estimates.

On the other hand, self-report has been found in other contexts to be a reliable and valid methodology for self-assessment of behavior. Based on their research employing both self-report and official data for 1,600 youth in Seattle, Hindelang, Hirschi, and Weiss (1981) showed that people's self-report of their behavior shows moderate to high consistency both within the same and across various self-report measures. Moreover, self-report data may offer more information regarding aggressive behavior than official school data such as referrals and suspensions, which do not reflect all instances of aggressive behavior. Self-report in principle should be able to identify delinquency even for people without criminal records (Hindelang et al., 1981). In short, the use of self-report may provide information not otherwise accessible to observers.

Apart from limitations in outcome measures, the present study is subject to much error, because it relied on potentially undependable sources: outcome evaluation methods

from the larger experimental evaluation and accounting expenditure data from the Anne Arundel County school district. The small number of schools implies large standard errors in the outcome evaluation. Self-selection also occurred: parental consent was received for only 3,167 out of 4,794 total students. Parents from low socioeconomic and ethnic minority backgrounds were less likely to grant consent for their children to participate. However, Harak's (2008) analysis of attrition with respect to student participation in the self-report survey did not find differential attrition for treatment and control groups. Lastly, the school district's accounting expenditure data may not encapsulate all costs incurred by the treatment and control interventions.

Because the groups were randomly assigned, randomization controlled for most threats to internal validity. Threats to external validity exist, however, because results may not generalize to other school districts such as urban and rural districts, since the Anne Arundel Public County School System is a relatively diverse suburban school district. Schools that participated in the outcome study are also mostly located in affluent areas, so results may not generalize to schools located in less affluent areas. Furthermore, the present study only examines one social and character development program—it is possible that results would be very different with other models or variations of the program. For example, this particular implementation of Second Step incorporated research components such as implementation logs and feedback, which are normally not a part of the Second Step program. These additional research components might affect the results by increasing the integrity of program implementation and the amount of time teachers spent in social and character development activities, as well as costs of research staff time spent in implementation.

Virtues of the Research

Despite its limitations, this study offers some insight on the costs and cost-effectiveness of Second Step in a suburban school system, including the salary and benefits of Second Step personnel and their time use. Results offer some evidence that this intervention is no more cost-effective than the alternative of business as usual. Members of the scientific community may be able to share these findings with school administrators and educators to consider in policy decisions regarding whether or not to expand the Second Step intervention to additional elementary schools within the school district.

Implications for Research and Practice

Since the school district is continuing to implement Second Step in the same schools following the three-year experimental period, this study provides valuable information about the costs of starting Second Step in additional schools. With insight about the average marginal costs involved in operating an enhanced whole-school intervention, school administrators may be able to make more informed decisions about resource allocation for the intervention.

Second Step, even under an enhanced implementation, does not improve cost-effectiveness over business as usual according to the present results. With a slightly lower marginal cost per student than the control programs after a three-year implementation period, the enhanced Second Step intervention will not ameliorate student outcomes any more than the control programs. However, this cost is sensitive to variations in teacher salaries and benefits and the number of students enrolled. Hence, unless Second Step can be shown to be effective at improving student outcomes, implementing Second Step in

additional schools cannot be justified on the grounds of cost-effectiveness. Instead, school resources should be directed towards exploration and implementation of programs that have been shown to be efficacious—or other plausible programs whose efficacy is under evaluation. As the literature review points out, few programs have been shown to be efficacious, so it is important that school districts allocate funds to those that are.

The results do not justify strong recommendations regarding Second Step one way or the other, as the program appears no more or less costly and no more or less efficacious than usual programming in these schools. Nevertheless, the non-trivial startup costs provide a reason to forego starting the program where it does not exist. Therefore, based on results, it is recommended that the Second Step intervention not be implemented in additional elementary schools in the school district. It is highly likely that these schools have existing social and character development programs; hence, it is recommended that these programs be evaluated for cost-effectiveness or be replaced with programs that are considered cost-effective. Although Second Step is no more costly and no more efficacious than the existing programs being implemented, the school district would incur a substantial startup cost in implementing Second Step. For the first year of implementation, startup costs are $438,683. Schools without the program are presumably funding existing social and character development programs, so these schools would not sustain new startup costs. Without evidence about the efficacy of Second Step, the large startup cost cannot be justified. Instead of starting Second Step in additional schools, the school district might set aside their funds to implement programs shown to be cost-effective.

Apart from its practical implications, this study provides insight about limitations in the typical methodology for analyzing costs. In the traditional ingredients approach, each resource required to implement the intervention is identified, and a cost value is placed on each separately. Costs of resources are typically determined by reviewing information from accounting expenditures and budgets, which are too often inaccurate and unreliable. Also, under the ingredients approach, overlap in costs exists between shared resources, rendering it challenging to separate costs that are devoted primarily to the intervention. Assumptions, such as the amount of or percentage of time used, are required to determine which portions of shared resources are utilized for the intervention.

Another flaw with the ingredients approach is difficulty assessing whether or not a resource should warrant inclusion in total cost estimates. The cost analysis relies on judgment about which resources are required for program implementation and should be included. Thus, the ingredients approach for measuring costs of programs and interventions requires speculation, assumption, and judgment. As a result, it likely that much error may occur in estimates of total marginal costs: costs of resources necessary to the intervention may be overlooked or incorrectly calculated. Further research may be warranted to explore the use of cost analysis methodologies with potential to reduce error.

To conclude, high-quality cost-effectiveness analyses of social skills and behavior programs are scarce. The school psychology field would benefit from further cost-effectiveness analyses of school-based preventative and early intervention educational programs in other school districts, once these programs have demonstrated efficacy. Comparisons may consider differences in results across various school districts' implementation of Second Step and may use alternative cost-effectiveness methodologies.

Appendix A. Comparison of Treatment and Control School Activities and Programs
Shortly after the Beginning of Intervention Implementation

Category	Treatment	Control
Violence prevention / peace promotion activities	conflict resolution Core Essentials county code of conduct guidance lessons review school rules and policies informal discussions Second Step	conflict resolution Core Essentials county code of conduct guidance lessons review rules / expected behavior PBIS no bullying peer mediation
Social / emotional development Activities	Core Essentials guidance lessons Second Step	Core Essentials guidance lessons "Don't Laugh at Me" program caught being good coupons classroom rules and behavior cooperative discipline discussions as problems arise respect mini-lessons teams to promote friendships and working together -- winning team recognized in newsletter
Character education activities	Character Counts Core Essentials county code of conduct guidance lessons monthly character trait student of the week children assess situations as a class discuss character traits of characters and how they parallel real people good citizenship club Second Step	Character Counts Core Essentials county code of conduct guidance lessons monthly character trait character tea star student behavior flip card system student teams receive points as they achieve desirable outcomes
Tolerance / diversity activities	Core Essentials cultural diversity day guidance lessons one-on-one conversations with students open court teammates story Social studies parent volunteer taught sign language Second Step	Core Essentials share students heritage projects guidance lessons class discussions on acceptance of everyone open court teammates story Social studies Language arts unit on friendship "Don't Laugh at Me" program
Risk prevention / health promotion activities	county code of conduct drug education with school nurse science health unit class discussions Core Essentials red ribbon week Second Step	county code of conduct nurse-directed health activity science health unit guidance lessons "Keep a Clear Mind" program

Appendix A. Comparison of Treatment and Control School Activities and Programs
Shortly after the Beginning of Intervention Implementation

Civic responsibility / community service activities	bake sales for local charities clothing / school supplies drive food drives / toy drives Human relations committee packages for soldiers Social studies buddy program career day student government	charity collections collecting / delivering items for women's shelter, cooking and serving food to women food drives / toy drives Human relations committee writing letters to soldiers in Iraq Social studies raising money for school in Kenya recycling
Behavioral management programs used in class	clip system class reward (parties) compliment tallies daily behavior charts drawings for prizes school-wide flip chart system incentive charts tickets for positive behavior team tally marks for good behavior points awarded for positive behavior and deducted for negative behavior rewards from prize box sticker charts super table reward super student bucket student of the day / week good behavior notes home weekly/daily progress reports weekly / monthly rewards character chart of positive traits good citizen award grab bag names listed on board for smile club Second Step	clip system class reward (parties) compliment points daily behavior charts drawing for prizes flip cards incentive charts tickets for positive behavior team tally marks for good behavior points awarded for positive behavior and deducted for negative behavior rewards from prize box sticker charts star tables mystery student student of the day / week good behavior notes home parent communication weekly/ monthly rewards good worker slips homework stars team of the week Core Essentials

Note. These data are based on a survey conducted shortly after the intervention began and do not represent a true baseline.

Appendix B. Cost studies included in the literature summary

Author	Program	Outcome Evaluation Design	Sample Size	Comparison Group	Cost Study	Assessment of Credibility
Masse and Barnett (2002)	California Abecedarian Program	Experimental; used only one IQ measure at each age	50 treatment students at age three	48 control students at age three	Cost-benefit analysis; gross exaggeration of benefits	Low
Schweinhart (2003)	High/Scope Perry Program	Experimental, but inadequate random assignment	58 treatment students	65 control students	Cost-benefit analysis; gross exaggeration of benefits	Low
Brewer, Krop, and Reichardt (1999)	National Class Size Reduction	Non-experimental	N/A	N/A	Cost analysis; no consideration of opportunity costs	Low
Krueger (2003)	Project STAR	Experimental; inner city schools may have been oversampled	17 treatment schools	22 control schools	Cost-benefit analysis; exaggeration of benefits	Low
Wasik and Slavin (1993)	Reading Recovery	"Best evidence" synthesis; includes qualitative studies that do not report effect sizes; no clear criteria for inclusion	N/A	N/A	N/A	Low
Marina and Gilman (2003)	Reading Recovery	Quasi-experimental; students who made no progress in the program were not counted as participants	95 treatment students	Non-equivalent sample of 99 control students	Cost-benefit analysis; exaggeration of benefits	Low
Grossen, Coulter, and Ruggles (1997)	Reading Recovery	Non-experimental	N/A	N/A	Cost-benefit analysis	Medium

Appendix B. Cost studies included in the literature summary

Madden, Slavin, Karweit, Dolan, and Wasik (1993)	Success for All (SFA)	Quasi-experimental, but schools not adequately matched	Five treatment schools	Five non-equivalent control schools	Cost analysis; no consideration of opportunity costs; exaggeration of benefits	Low
Borman and Hewes (2002)	Success for All (SFA)	Quasi-experimental, but schools not adequately matched	Five treatment schools	Five non-equivalent control schools	Cost-effectiveness analysis; inconsistent discounting	Low
King (1994)	Success for All (SFA), Accelerated Schools, and School Development Program (SDP)	Relied on non-experimental evaluations of others	N/A	N/A	Cost analysis; used cost estimates of SFA and Accelerated Schools to estimate SDP costs	Medium
Traub (1999)	Multiple programs	Non-experimental; relied on AIR estimates	N/A	N/A	Only included implementation costs	Low
Haynes, Comer, and Hamilton-Lee (1988)	School Development Program (SDP)	Pre-experimental	Four treatment schools	None	None	Low
Hartman and Fay (1996)	Instructional Support Teams (IST)	Pre-experimental	1074 treatment schools	None	Cost-effectiveness analysis; range of discount rates inappropriate	Low
Scott and Barrett (2004)	Positive Behavior Support (PBS)	Case study; only one outcome measure	One treatment school	N/A	Cost analysis; only included training costs	Low
Foster and Jones (2005)	Fast Track	Experimental; standard errors large; only two outcomes measures	445 treatment students	446 control students	Cost-benefit analysis, based on limited budget information	Low
Embry (2002)	Good Behavior Game (GBG)	Pre-experimental; unknown methodology	15 treatment schools	None	Cost-effectiveness analysis; cost estimates based on a hypothetical implementation; exaggeration of benefits	Low
Aos, Lieb, Mayfield, Miller and Pennucci (2004)	Multiple programs	Meta-analysis	N/A	N/A	Cost-benefit analysis; failed to include personnel time and costs of training	Medium

Appendix C. Personnel Time Use Questionnaires:
Counselor, Principal, Program Staff and Research Staff

How Counselors Spend Time

This questionnaire is part of the cost evaluation of character education activities in our schools. We ask you to estimate the amount of time you spend in various activities. Your responses will be anonymous. We do not want your name on this form. The information will be used for research and statistical purposes only, and no information individually identifiable to you will ever be reported. Averages for groups of counselors will be used to help estimate the costs of character education activities in schools. Your help with this questionnaire is up to you. You have the right not to participate or not to answer any and all questions. While your participation is voluntary, your help is important to the success of the study. Completing this questionnaire will not benefit you personally; instead your participation will help us learn more about the costs of educational programs. Please answer the following questions, place the completed questionnaire in the envelope provided, and return it to Debbie Wooleyhand, Student Services Building. Use a pencil so that you can erase an answer you wish to change.

1. First, please estimate the total number of **hours** per typical week you spend on all activities related to your job as a counselor in your school.

_____ Hours per week (this number should equal the number in the bottom right cell in the table below)

2. Using the table below, please estimate how these hours are distributed across the activities listed. For each cell in the table, indicate the number of hours (or fractions of hours) spent in each type of activity. If you do not typically spend time in an activity, enter zero (0) in the cell for that activity.

Activity	Best estimate of hours per average or typical week in:			
	School-Wide Social and Character Development Activity	Classroom Social and Character Instructional Programs	All other counseling and other work activities	Total
Preparation during or outside of school hours, working alone				
Implementation monitoring and evaluation of school programs				
School record keeping/ paperwork (e.g. IEP minutes, program notes, etc.)				
Attending or facilitating meetings				
Providing training or consultation to teachers or administrators				
Ordering, creating, and distributing instructional or planning materials to teachers				
Individual or group counseling				
Interacting with parents outside of meetings				
Interacting with educators or administrators outside of the already listed activities				
Other activities not included above (e.g., standardized testing, surveys, fire drills, etc.)				
Total hours spent in all activities (add the hours in each column)				

How Principals Spend Time

This questionnaire is part of the cost evaluation of character education activities in our schools. We ask you to estimate the amount of time you spend in various activities. Your responses will be anonymous. We do not want your name on this form. The information will be used for research and statistical purposes only, and no information individually identifiable to you will ever be reported. Averages for groups of principals will be used to help estimate the costs of educational activities in schools. Your help with this questionnaire is up to you. You have the right not to participate or not to answer any and all questions. While participation is voluntary, your help is important to the success of the study. Completing this questionnaire will not benefit you personally; instead your participation will help us learn more about the costs of educational programs. Please answer the following questions, place in the envelope provided, and return to Debbie Wooleyhand, Student Services Building. Use a pencil so that you can erase an answer you wish to change.

1. Please estimate the total number of **hours** per typical week you spend on all activities related to your job as an administrator in your school.

_____ Hours per week (this number should equal the number in the bottom right cell in the table below)

2. Using the table below, please estimate how these hours are distributed across the activities listed. For each cell in the table, indicate the number of hours (or fractions of hours) spent in each type of activity. If you do not typically spend time in an activity, enter zero (0) in the cell for that activity.

Activity	Best estimate of hours per average or typical week in:			
	School-Wide Social and Character Development Activity	Classroom Social and Character Instructional Programs	All other administrative and other work activities	Total
Preparation during or outside of school hours, working alone				
Implementation monitoring and evaluation of school programs				
Record keeping/documentation: administrative, planning or accountability purposes				
Attending or facilitating meetings				
Providing training or consultation for school personnel				
Ordering, creating, and distributing instructional or planning materials to teachers				
Supervision and observation				
Planning and problem solving outside of the already listed activities				
Other administrative activity outside of the already listed activities				
Activities not included above (e.g., standardized testing, surveys, fire drills)				
Total time spent in all activities (add the hours in each column)				

Anne Arundel County Public School System Personnel Activities

This questionnaire is part of the cost evaluation of character education activities. We ask you to estimate the amount of time you spend in various activities related to research and program implementation. Your responses will be anonymous. The information will be used for research and statistical purposes only, and no information individually identifiable to you will ever be reported. Your help with this questionnaire is up to you. You have the right not to participate or not to answer any and all questions. While participation is voluntary, your help is important to the success of the study. Completing this questionnaire will not benefit you personally; instead your participation will help us learn more about the costs of educational programs. Please answer the following questions and return to Sharon Huang, 3214 Benjamin Building, College Park, MD 20742. Use a pencil so that you can erase an answer you wish to change.

Please estimate the total amount of time (in hours) involved in work you have completed relating to program implementation. For ongoing activities, please estimate costs incurred over one year. If you have not completed the activity, write N/A for the entire row.

Activity	Research	Second Step Implementation	Implementation of Other Character Ed	Other	Control
Record keeping (e.g., consent forms, SASI files, MPR, stipends, etc.)					
Ordering materials (e.g. behavioral observations, incentives, Second Step materials , character ed materials					
Distributing materials (e.g. surveys, logs, consent forms, Second Step kits, character ed materials					
Collecting surveys					
Collecting implementation logs					
In-school support during survey administration					
Meetings with counselors (both prep time and actual meeting time)					
Meetings with school systems personnel (e.g. strategies for getting back research materials, strategies for participating)					
Communication with counselors (e.g. going into schools to discuss research or implementation)					
Training support					
Conference calls with MPR in preparation for and during data collection					
Attending trainings (e.g. SACD conference with IES, problem-solving—force field analysis)					
Locating resources for character education					
Making materials (e.g. cups and strings, certificates, incentives, 2nd Step homework)					
Other – please list:					

Position:_____

University of Maryland Research Team Activities

This questionnaire is part of the cost evaluation of character education activities. We ask you to estimate the amount of time you spend in various activities related to program implementation. Your responses will be anonymous. The information will be used for research and statistical purposes only, and no information individually identifiable to you will ever be reported. Your help with this questionnaire is up to you. You have the right not to participate or not to answer any and all questions. While participation is voluntary, your help is important to the success of the study. Completing this questionnaire will not benefit you personally; instead your participation will help us learn more about the costs of educational programs. Please answer the following questions and return to Sharon Huang, 3214 Benjamin Building (CAPS). Use a pencil so that you can erase an answer you wish to change. Please estimate the total amount of time (in hours), supplies (a comprehensive list of those used), transportation costs (number of round-trips made) involved in work you have completed relating to program implementation in the year 2006-2007. For ongoing activities, please estimate costs incurred over *one year*. If you have not completed the activity, write N/A for the entire row.

Activity:	TIME	SUPPLIES	TRANSPORTATION
Setting implementation goals and standards			
Printing implementation logs			
Packing and shipping logs to the schools			
Replacement of missing/lost logs			
Scanning the implementation logs			
Creating implementation feedback			
Printing feedback			
Meeting time spent providing implementation feedback			
Summer meeting for School Improvement Team plans			
Other - please list:			

Appendix D. List of Research and Implementation Activities

Activity:
1- Record keeping
   a. All research (SASI, MPR, consent, stipends, etc.)
2- Ordering materials
   a. Research - for behavioral observations, for incentives to participate
   b. Implementation - 2nd Step materials, character education materials
3- Distributing materials
   a. Research - surveys, logs, consent forms
   b. Implementation – 2nd Step kits, character education materials
4- Collecting surveys
   a. Only research
5- Collecting implementation logs
   a. Only implementation
6- In-school support during survey administration
   a. Only research
7- Meetings with counselors (prep and actual time)
   a. Research - data collection, MPR issues, consent forms, preparation of incentives
      for participating in surveys or getting back research forms, etc.
   b. Implementation – 2nd Step feedback, sharing of character education tips,
      preparation of character education ideas or incentives for 2nd Step
      implementation or returning logs (bookmarks, etc.)
8- Meetings with school systems personnel
   a. Research – strategies for getting back research materials, administration buy-in
      about data collection
   b. Implementation - strategies for getting: counselors to meetings, staff to summer
      meetings and trainings, implementation logs returned, 2nd Step scheduled into the
      school week.
9- Communication with counselors
   a. Going in to schools to discuss either research or implementation (see above for
      examples)
10- Training support
   a. Implementation – could be 2nd Step-related or for other character development
11- Conference calls
   a. Only research
12- Attending trainings
   a. Research - this could be the SACD conference with IES, or any other trainings
      you may have attended related to our evaluation or evaluation in general.
   b. Implementation - can be trainings for 2nd Step or other related programs (PBIS,
      etc) or problem-solving (Gary's force field analysis)
13- Locating resources for character education
14- Making materials
   a. Research – behavior observation prep (cups and strings), certificates and
      incentives for completing research forms
   b. Implementation – working on 2nd Step homework, incentives for implementation
15- Other duties not related to Second Step

Appendix E. Implementation Log Example
Grade 3 Second Step Implementation Record
Unit I: Empathy Training

Instructions: Please mark ⓪ if you did not complete that portion of the lesson. Mark ① if you completed the portion of the lesson, but with some modifications. Mark ② if you completed the portion of the lesson exactly as prescribed in the Second Step lesson materials.

Date Lesson Delivered   **Lesson 1: Empathy Training-Skill Overview**

Month	Day		In conducting this lesson, did you . . .
◯ Sep	⓪ ⓪	⓪ ① ②	1. Distribute take-home letter 1: Introduction to *Second Step*?
◯ Oct	① ①	⓪ ① ②	2. Introduce what will be learned in using *Second Step*?
◯ Nov	② ②	⓪ ① ②	3. Instruct students what to look for to discern what people feel?
◯ Dec	③ ③	⓪ ① ②	4. Play the video, *Lunchtime*?
◯ Jan	④ ④	⓪ ① ②	5. Show Photograph A?
◯ Feb	⑤ ⑤	⓪ ① ②	6. Discuss how Jenni feels based on the clues?
◯ Mar	⑥ ⑥	⓪ ① ②	7. Explain that people can feel several feelings at once?
◯ Apr	⑦ ⑦	⓪ ① ②	8. Discuss how Meg and Brian feel based on the clues?
◯ May	⑧ ⑧	⓪ ① ②	9. Explain that everyone has different feelings about the situation?
◯ Jun	⑨ ⑨	⓪ ① ②	10. Discuss how Kyle's actions may affect Jenni?
		⓪ ① ②	11. Show photograph B while showing second part of video?
		⓪ ① ②	12. Discuss how Kyle and Jenni feel after second part of video?
		⓪ ① ②	13. Discuss how feelings change?
		⓪ ① ②	14. Explain that Kyle's actions were an accident?
		⓪ ① ②	15. Model acting out a scenario?
		⓪ ① ②	16. Have students discuss the feelings you acted out?
		⓪ ① ②	17. Students role play scenarios for the class to identify feelings?

⓪	①	②	③	④	18. How many minutes did you spend on this lesson?
< 20 min.	21-30 min.	31-40 min.	41-50 min.	> 50 min.	

Comments about lesson 1:

Appendix F. Second Step Transfer of Training Checklist Example

Second Step Transfer of Training Checklist: Unit III: Anger Management

Please fill in the appropriate box to indicate how frequently the following events occurred in the **past week** *OUTSIDE* of the Second Step lesson.

How often did you...	Never	One or two times	Every day	Several times each day
1. Use Imagine the Day activities to help students identify and target times during the day when they might use new Second Step skills on their own?	Ⓐ	Ⓑ	Ⓒ	Ⓓ
2. Provide coaching and prompting to students about when to use specific Second Step skills?	Ⓐ	Ⓑ	Ⓒ	Ⓓ
3. Use "thinking out loud" to model perspective-taking, problem-solving, and/or anger-management steps?	Ⓐ	Ⓑ	Ⓒ	Ⓓ
4. Use verbal praise and encouragement to reinforce student use of new Second Step skills (i.e. calming down strategies, evaluating solutions to problems)?	Ⓐ	Ⓑ	Ⓒ	Ⓓ
5. Use Remember the Day activities at the end of the day to encourage students to share how they used Second Step skills throughout the day?	Ⓐ	Ⓑ	Ⓒ	Ⓓ
6. Use Second Step language (i.e. empathy, anger buttons) in my interactions with students during non-Second Step activities?	Ⓐ	Ⓑ	Ⓒ	Ⓓ
7. Reinforce students for using Second Step language by prompting students to name the skill and describe the benefits of using it?	Ⓐ	Ⓑ	Ⓒ	Ⓓ
8. Intervene in student conflict(s) by prompting students to use anger-management or social-problem solving strategies (i.e. had students share how the other was feeling, brainstorm solutions)?	Ⓐ	Ⓑ	Ⓒ	Ⓓ
9. Prompt students to assess whether they might need to use calming down strategies to manage their emotions?	Ⓐ	Ⓑ	Ⓒ	Ⓓ
10. Praise students when I noticed displays of pro-social behaviors towards others (i.e. helping, sharing, active listening)?	Ⓐ	Ⓑ	Ⓒ	Ⓓ

Additional Comments:

References

Abramson, P. (2008). *More dollars spent, less construction completed* (Supplement to
    School Planning and Management). Retrieved March 20, 2008 from the National
    Clearinghouse for Educational Facilities Web site:
    http://www.edfacilities.org/ds/index.cfm

Aos, S., Lieb, R., Mayfield, J., Miller, M., Pennucci, A. (2004). *Benefits and costs of
    prevention and early intervention programs for youth.* Olympia, WA: Washington
    State Institute for Public Policy.

Aos, S., Phipps, P., Barnoski, R., & Lieb, R. (2001). *The comparative costs and benefits
    of programs to reduce crime.* Olympia, WA: Washington State Institute for Public
    Policy.

Barnett, W.S. (1985). Benefit-cost analysis of the Perry Preschool Program and its policy
    implications. *Educational Evaluation and Policy Analysis, 7*, 333-342.

Barnett, W.S. (1993). Benefit-cost analysis of preschool education: Findings from a 25-
    year follow-up. *American Journal of Orthopsychiatry, 63,* 500-508.

Borman, G.D., & Hewes, G. (2002). The long-term effects and cost-effectiveness of
    Success for All. *Educational Evaluation and Policy Analysis, 24*, 243-266.

Brewer, D. J., Krop, C., Gill, B. P., & Reichardt, R. (1999). Estimating the cost of
    national class size reductions under different policy alternatives. *Educational
    Evaluation and Policy Analysis, 21*, 179-192.

Carr, E. G., Horner, R. H., Turnbull, A. P., Marquis, J. G., Magito McLaughlin, D.,
    McAtee, M. L., Smith, C. E., Anderson Ryan, K., Ruef, M. B., & Doolabh, A.
    (1999). *Positive behavior support for people with developmental disabilities: A
    research synthesis.* Washington, DC: American Association on Mental
    Retardation.

The David and Lucile Packard Foundation. (2007). Appendix A: School facilities. *The
    Future of Children, 7,* 127-132. Retrieved March 20, 2008 from the Future of
    Children Web site: http://www.futureofchildren.org

Committee for Children. (2006). Frequently Asked Questions. Retrieved July 31, 2007
    from the World Wide Web: http://www.cfchildren.org/cfc/aboutf/faq

Eitel, S.B., Lamberth, J.J., & Hyman, I.A. (1984). Time utilization study of school
    psychologists in an urban setting. *Psychology in the Schools, 21*, 329-335.

Fletcher, J.D., Hawley, D.E., & Piele, P.K. (1990). Costs, effects, and utility of

microcomputer assisted instruction in the classroom. *American Educational Research Journal, 27*, 783-806.

Foster, E.M., Jones, D.E., & the Conduct Problems Prevention Research Group. (2006). Can a costly intervention be cost-effective? An analysis of violence prevention. *Archives of General Psychiatry*, *63,* 1284-1291.

Foster, E.M., Shelton-Johnson, D., Taylor, T. (2007). Measuring time costs in psychosocial interventions designed to reduce behavior problems among children and youth. *American Journal of Community Psychology, 40,* 64-81.

Frey, K. S., Hirschstein, M. K., & Guzzo, B. A. (2000). *Second Step*: Preventing aggression by promoting social competence. *Journal of Emotional and Behavioral Disorders, 8*, 102-112.

Frey, K. S., Nolen, S.B., Edstrom, L.V.S., & Hirschstein, M.K. (2005). Effects of a school-based social-emotional competence program: Linking children's goals, attributions, and behavior. *Journal of Applied Developmental Psychology, 26*, 171-200. http://www.promisingpractices.net/program.asp?programid=111

Gottfredson, G.D. (2006). *Character education inquiry: Summary* (Year 3 Progress Rep.). College Park, MD: University of Maryland.

Gottfredson, G.D. (2005). Hawthorne effect. In B. Everitt & D. Howell (Eds.), *Encyclopedia of statistics in behavioural science.* Chichester, UK: Wiley.

Gottfredson, G.D., & Gottfredson, D.C. (2001). What schools do to prevent problem behavior and promote safe environments. *Journal of Educational and Psychological Consultation, 12*, 313-344.

Gottfredson, G.D., Harak, E.T., Nebbergall, A., Nese, J., & Shaw, F. (2008, June). *An experimental evaluation of an elementary school social competency instruction program.* Paper presented at the Social and Character Development Research Program Consortium Meeting, Arlington, VA.

Greenlee, B.J. & Bruner, D.Y. (2001). Effects of schoolwide reform strategies on Title I schools using Success for All compared to Title I schools using other literacy enhancing strategies. *Education, 122*, 177-188.

Grossman, D.C., Neckerman, H.J., Koepsell, T.D., Liu, P.Y., Asher, K.N., Beland, K., Frey, K., & Rivara, F.P. (1997). Effectiveness of a violence prevention curriculum among children in elementary school: A randomized controlled trial. *Journal of the American Medical Association, 277,* 1605-1611.

Grossen, B., Coulter, G. A., & Ruggles, B. (1997). Reading Recovery: An evaluation of benefits and costs. *Effective School Practices 15*, 6-24.

Hanushek, E.A. (1999). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Education Evaluation & Policy Analysis, 21*, 143–163.

Harak, E.T. (2008). *Outcomes of an elementary grades social competence experiment according to student self-report*. Unpublished doctoral dissertation. University of Maryland, College Park.

Hartman, W.T. & Fay, T.A. (1996). Cost-effectiveness of instructional support teams in Pennsylvania (Policy Paper No. 9). Palo Alto, CA: Center for Special Education Finance, American Institutes for Research.

Haynes, N. M., Comer, J. P., and M. Hamilton-Lee (1988). The School Development Program: A model for school improvement. *Journal of Negro Education, 57,* 11-21.

Hindelang, M. J., Hirschi, T., & Weiss, J.G. (1981). *Measuring delinquency*. Newbury Park, CA: Sage.

Hülsmann, T. (2007). Costing Open and Distance Learning. Retrieved November 27, 2007 from the World Wide Web: http://www.col.org/trainingresources/costingodl/

Jones, E.M., Gottfredson, G.D., Gottfredson, D.C. (1997). Success for some: An evaluation of a Success For All program. *Evaluation Review, 21*, 643-670.

Juster, F.T. (1985). Conceptual and methodological issues involved in the measurement of time use. In F.T. Juster & F.P. Stafford (Eds.), *Time, goods, and well-being* (pp. 19-31). Ann Arbor, MI: University of Michigan, Institute of Social Research.

Juster, F.T. & Stafford, F.P. (1991). The allocation of time: Empirical findings, behavioral models and problems of measurement. *Journal of Economic Literature, 29*, 471-522.

Kavale, K.A. & Forness, S.R. (1999). *Efficacy of special education and related services*. Washington, DC: American Association of Mental Retardation

King, J.A. (1994). Meeting the educational needs of at-risk students: A cost analysis of three models. *Educational Evaluation and Policy Analysis*, *16*, 1-19.

Krueger, A.B. (2003), Economic considerations and class size. *The Economic Journal, 113,* 34-63.

Levin, H.M. (1983). *Cost-effectiveness: A primer*. Newbury Park, CA: Sage.

Levin, H.M. (1987). Accelerated Schools for disadvantaged students. *Educational

Leadership, 44, 19-21.

Levin, H.M. & McEwan, P.J. (2001). *Cost effectiveness analysis: Methods and applications*. Thousand Oaks, CA: Sage.

Levin, H.M., Glass, G.V., & Meister, G.R. (1987). Cost effectiveness of computer-assisted instruction. *Evaluation Review, 11,* 50-72.

Madden, N.A., Slavin, R.E., Karweit, N.L. Dolan, L.J., & Wasik, B.A. (1993). Success for All: Longitudinal effects of a restructuring program for inner-city elementary schools. *American Educational Research Journal, 30*, 123-148.

Marina, B. & Gilman, D.A. (2003). Is reading recovery worth the cost? (ERIC Document Reproduction Service No. ED 473 957)

Masse, L. N., & Barnett, W. S. (2002). *A benefit-cost analysis of the Abecedarian Early Childhood Intervention*. New Brunswick, NJ: National Institute for Early Education Research, Rutgers University.

Niemiec, R.P., Sikorski, M.F., & Walberg, H.J. (1989). Comparing the cost-effectiveness of tutoring and computer-based instruction. *Journal of Educational Computing Research, 5,* 395-407.

Robinson, J.P. & Godbey, G. (1999). *Time for life: The surprising ways Americans use their time* (2nd ed.). University Park, PA: The Pennsylvania State University Press.

Ross, J.A., Barkaoui, K., & Scott, G. (2007). Evaluations that consider the cost of educational programs. *American Journal of Evaluation, 28,* 477-492.

Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A systematic approach.* Thousand Oaks, CA: Sage Publications.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*, 688-701.

Schweinhart, L.J. (2003, April). *Benefits, costs, and explanation of the High/Scope Perry Preschool Program*. Paper presented at the 2004 biennial meeting of the Society for Research in Child Development, Tampa, FL.

Scott, T.M., & Barrett, S.B. (2004). Using staff and student time engaged in disciplinary procedures to evaluate the impact of school-wide PBS. *Journal of Positive Behavior Interventions, 6(1)*, 21-27.

Slavin, R.E., Dolan, L., Madden, N. A., Karweit, N. L., & Wasik, B.A. (1992). Success for All policy implications (CDS Report No. 35). Baltimore, MD: Johns Hopkins University, Center for Social Organization of Schools.

Society for Prevention Research. (2004). Standard of Evidence. Retrieved January 7, 2008 from the World Wide Web: http://www.preventionresearch.org/StandardsofEvidencebook.pdf292.5 k

Tanner, C.K. (2000). Minimum Classroom Size and Number of Students per Classroom. Retrieved March 20, 2008 from the World Wide Web: http://www.coe.uga.edu/sdpl/research/territoriality.html

Traub, J. (1999). *Better by design? A consumer's guide to schoolwide reform.* Washington, DC: Thomas Fordham Foundation.

U.S. Department of Education. Safe, Disciplined and Drug-Free Expert Panel. (2001). Exemplary Programs 2001. Retrieved July 31, 2007, from the World Wide Web: http://www.ed.gov/admins/lead/safety/exemplary01/panel_pg2.html

Walker, D., & Kumaranayake, L. (2002). How to do (or not to do) . . . allowing for differential timing in cost analyses: discounting and annualization. *Health Policy and Planning, 17*, 112-118.

Wasik, B.A., & Slavin, R.E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly 28*, 178-200.